NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

USING A FUNCTIONAL ARCHITECTURE TO IDENTIFY HUMAN-AUTOMATION TRUST NEEDS AND DESIGN REQUIREMENTS

by

Bradley A. Johnson

December 2016

Thesis Advisor: Joseph W. Sweeney III
Second Reader: Karen S. Holness

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 2016	**3. REPORT TYPE AND DATES COVERED** Master's Thesis 01-05-2015 to 12-16-16

4. TITLE AND SUBTITLE
USING A FUNCTIONAL ARCHITECTURE TO IDENTIFY HUMAN-AUTOMATION TRUST NEEDS AND DESIGN REQUIREMENTS

5. FUNDING NUMBERS

6. AUTHOR(S) Bradley A. Johnson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB number ____N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

This thesis develops and analyzes the functional architecture for an "autonomous" unmanned aerial system performing an Intelligence, Surveillance, and Reconnaissance (ISR) mission without a continuous communication link to human operators for trust needs. The factors that affect human trust are developed from a literature review covering theory and empirical studies that have investigated the importance of human trust in human-automation interactions. The identified factors are applied to the functional architecture, and the system functions are categorized as Reasoning functions and Non-reasoning functions. Each functional category is analyzed for trust needs by describing how the function's purpose, process, and performance link to human knowledge, perception and beliefs. From the analysis, automation design requirements that link to the identified trust needs are developed. This work highlights the importance of applying human factors analyses in the early stages of the Systems Engineering process for "autonomous" systems.

14. SUBJECT TERMS automation, autonomy, autonomous, trust, human systems integration, human factors, systems engineering, systems engineering process	**15. NUMBER OF PAGES** 103
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

USING A FUNCTIONAL ARCHITECTURE TO IDENTIFY HUMAN-AUTOMATION TRUST NEEDS AND DESIGN REQUIREMENTS

Bradley A. Johnson
Lieutenant, United States Navy
B.S., Iowa State University, 2008

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
December 2016

Approved by: Joseph W. Sweeney III
Thesis Advisor

Karen S. Holness, Ph.D.
Second Reader

Ronald Giachetti, Ph.D.
Chair, Department of Systems Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis develops and analyzes the functional architecture for an “autonomous” unmanned aerial system performing an Intelligence, Surveillance, and Reconnaissance (ISR) mission without a continuous communication link to human operators for trust needs. The factors that affect human trust are developed from a literature review covering theory and empirical studies that have investigated the importance of human trust in human-automation interactions. The identified factors are applied to the functional architecture, and the system functions are categorized as Reasoning functions and Non-reasoning functions. Each functional category is analyzed for trust needs by describing how the function’s purpose, process, and performance link to human knowledge, perception, and beliefs. From the analysis, automation design requirements that link to the identified trust needs are developed. This work highlights the importance of applying human factors analyses in the early stages of the Systems Engineering process for “autonomous” systems.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AA	autonomous/automated agent
ATEVV	Autonomy Community of Interest Test and Evaluation, Verification and Validation
AVACC	aerial vehicle with automated cognitive capability
CCOI	critical contact of interest
CPA	closest point of approach
COI	contact of interest
DOD	Department of Defense
DSB	Defense Science Board
EM	electromagnetic
ESM	electronic support measure
F2T2EA	find, fix, track, target, engage and assess
FFBD	functional flow block diagram
GPS	Global Positioning System
HFE	Human Factors Engineering
HVU	high value unit
IEEE	Institute of Electrical and Electronics Engineers
IR	infrared
ISR	intelligence, surveillance and reconnaissance
SE	Systems Engineering
UAV	unmanned aerial vehicle

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The driving factor for this research originates from a Defense Science Board (DSB) task report from 2012. The task report lists five research topics as recommended areas of investigation into the military's future use of "autonomous" systems. This thesis relates to two of the research topics. The first topic is the "methods and metrics for confirming that an autonomous system will perform or interact with its human as intended and for measuring the user's trust in the system" (DSB 2012, 12). The second topic is the "interfaces that make the basis of autonomous system decisions more apparent to its users" (DSB 2012, 12). This thesis concentrates on one DSB identified capability gap that is discussed under a category that is relevant to both of the research topics. The category is "understandable autonomous behaviors" (DSB 2012, 48), and the identified gap from the category is, "models of what operators or decision makers need to know about the system or state in order to maintain trust in the predictable outcomes from using the system" (DSB 2012, 49).

This research develops the first iteration of trust-centered design requirements for a hypothetical advanced "autonomous" aerial system. The requirements trace to identified needs for "understanding autonomous behaviors" (DSB 2012, 48) that can aid with calibrating and maintaining trust in the system. This is completed by applying the findings from multiple published empirical studies and accepted psychological theory to the Functional Analysis stage of the Systems Engineering process.

The human psychological factors that influence trust are described in the following definition which was developed from investigating multiple published research papers related trust in automation and the Theory of Reasoned Action:

> Trust is the attitude of a human, developed from beliefs, perceptions and knowledge of a system's functional capabilities, towards the behavior of reliance in the system's actions to achieve the human defined goals in situations characterized by uncertainty and vulnerability. (Ajzen 2007)

From the definition, trust influences behavior; and, human behavior is one of the many considerations Human Factors Engineering (HFE) practitioners investigate when

"designing for human use" (Lockett and Powers 2003, 463). Therefore, HFE considerations, and thus trust, are important to the system design. Methods for performing HFE analysis include task analysis and function allocation (Lockett and Powers 2003, 464). Thus, a functional architecture was developed to apply these methods.

The functional architecture developed for the hypothetical advance "autonomous" aerial system was based on a capability currently met by a manned aircraft system: identifying vessels on the open ocean during an Intelligence, Surveillance, and Reconnaissance (ISR) mission without communication with another system. Analysis performed on the hypothetical system focused on the functions in which the system must be given a high degree of autonomy to meet the desired capability.

The key finding from this research is that an analysis performed on the functional architecture of an "autonomous" system linking human psychological factors that influence trust to the system's functional purpose, process, and performance (Lee and See 2004) can aid in developing system design requirements that directly trace to trust calibration and correct reliance in the system. This analysis can, and should be performed early in the Systems Engineering Process (Blanchard and Fabrycky 2011, 33) before the detailed system design and detailed interface design are developed.

References

Ajzen, I. 2007. *Attitudes, Personality and Behavior*. McGraw-Hill Education. Accessed August 6, 2016. ProQuest Ebook Central. http://ebookcentral.proquest.com.libproxy.nps.edu/lib/ebook-nps/detail.action?docID=287791&fPQ=1

Blanchard, Benjamin S. and Wolter J Fabrycky. 2011. *Systems Engineering and Analysis.* 5th ed. Upper Saddle River: Prentice Hall.

Defense Science Board. 2012. *Task Force Report: The Role of Autonomy in DOD Systems*. Department of Defense, Defense Science Board Office of the Under Secretary of Defense for Acquisition, Technology and Logistics. July 2012. Washington, DC, http://www.acq.osd.mil/dsb/reports/AutonomyReport.pdf.

Lee, John D. and Katrina A. See. 2004. “Trust in Automation: Designing for Appropriate Reliance.” *Human Factors*, Vol. 46, No. 1, Spring: 50–80. http://dx.doi.org/10.1518/hfes.46.1.50_30392.

Lockett III, John F. and Jeffrey Powers. 2003. “Human Factors Engineering Methods and Tools.” In *Handbook of Human System Integration*, edited by Harold R. Booher, 463–496. New Jersey: John Wiley and Sons, Inc.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge the steady support, love, and understanding that my wife and children have provided during my graduate studies. Late evenings and short weekends can take a toll on family life and relationships, yet, somehow I feel like we are exiting this tour in Monterey closer than ever.

Secondly, I would like to extend my gratitude to my advisor Joe Sweeney and to my second reader Karen Holness. Thank you both for sharing your time and expertise which were both instrumental in my research and in my graduate studies in general.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

This work begins with a short scenario describing an existing military capability critical to battlespace awareness. The capability must be retained, but how it is met will change if the systems used include an increased number of automated processes than are currently in use.

A. MANNED VS. UNMANNED CAPABILITY GAP

The crew of an air-capable warship operating independently on the open ocean desires to extend awareness of the operational environment beyond its own sensor horizon to gain tactical advantage over a potential adversary. Typically, a manned helicopter such as the MH-60R is launched to fulfill this capability need. The helicopter's sensors include a multi-mode radar, optical and infrared cameras, passive electronic sensors, real-time data link, and a fully trained crew. The helicopter not only extends the ship's sensor horizon via the data link but also brings its own capabilities to the mission that extend beyond data link and voice communication ranges with the ship. The operating range of the helicopter allows for flying beyond the communication range, and the capability of the aircrew team using the helicopter sensor suite allow for collecting and interpreting data from the extended operating environment. When the aircrew is able, or desires to report findings beyond the communication range, they must either climb to a higher altitude or decrease their distance from the ship's position to regain communication. Once communication is reestablished, the collected information is transmitted by the aircrew to the ship via data link or voice. The ship can expect the aircrew to report information in a known format, and provide all relevant collected data points. If uncertainty exists in any of the data, the aircrew can verbalize this in their report to the ship. No further interpretation of the transmitted information is necessary by the ship's crew. The employment of the helicopter meets the ship crew needs and the sensor horizon is extended.

Adding unmanned aerial vehicles (UAVs) to the Navy's inventory of systems has allowed the described capability to be augmented. Instead of sending a manned helicopter

to extend the ship's sensor horizon, a UAV can be sent which removes the aircrew from any potential threat that may exist in the operational area and from the hazards of physically being in the aircraft. The traditional aircrew responsibilities still exist for the UAV, but many of functions normally performed in the air are changed. Instead of physically controlling the aircraft in flight, the pilot now remotely directs and monitors the UAV on a flight path. The sensor operator(s) functions are not any different from what they are in the air other than now sensors are tele-operated via a data link. Much like in the manned aircraft, the aircrew uses their training combined with the UAV sensor suite to interpret the environment they are experiencing through the UAV sensors. The most significant difference between using a manned aircraft and using a UAV to fill this capability is that a continuous communication link must exist between the aircrew and the UAV. While the UAV has removed the aircrew from potential threats in the operational area and also from the hazards that flying creates, the UAV, as currently used, lacks the same capabilities a manned aircraft has when operating beyond the limits of a communication link. This is the existing UAV capability gap. This research will investigate the design needs for an aerial system intended to operate without a continuous communication link.

B. OVERVIEW OF AUTOMATION USE

Automation has been used to supplement, and in some cases completely replace, a variety of human work processes. Sophisticated algorithms, speedier processing, and miniaturization of components together have allowed engineers to develop systems that can, and sometimes do perform processes or actions normally carried out by a human, faster and with fewer errors than a human ever could. When automated processes are used the human operator is replaced or moved to a different role in the process. The result is a system in which the automated machine performs functions traditionally performed by a human operator with the human filling the role of teammate, monitor, or supervisor rather than operator. Depending on the type of interactions a human has with the automated machine the human-machine system can be classified as a human-on-the-loop system or a human-in-the-loop system. The human-in-the-loop classification comes from descriptions of simulators used for testing and evaluating human-automation systems.

Human-in-the-loop simulators require humans to complete certain actions (Sheridan 2002, 131) before, or in response to automation actions for a functional process to continue. The human is therefore in the functional "loop." From this description, it can be reasoned that the human-on-the-loop classification describes a system in which a functional process does not require specific human actions to continue.

The human-in-the-loop and human-on-the-loop classifications are useful for describing general human-automation interaction in a system, but as systems become more complex, it is better to use these descriptors in terms of system capabilities. When the human is primarily acting as a monitor or supervisor, these integrated systems are usually described as human-on-the-loop systems. The human monitors the machine's performance and provides instruction when or if needed. When the human is filling the roll of teammate, and in some cases supervisor or monitor, the system is described as a human-in-the-loop system. In this case the machine depends on human input or approval to start, continue, or complete a process. Depending on the process the system is executing, the human can be both in-the-loop or on-the-loop for the same system.

As mentioned in Section A, one clear example of advanced use of automation in today's military is in the use of UAVs. Current UAV processes fit into both the in-the-loop and on-the-loop classifications. The pilot's role in the system fits both the monitor's and supervisor's role by entering flight commands and monitoring system responses. The UAV's onboard flight computer takes the flight commands and manipulates electric, hydraulic, mechanical, and power subsystems to alter the aerodynamics and thrust of the vehicle to achieved the desired flight plan. The pilot acts ***on-the-loop*** in the sense that aerodynamic and thrust control of the vehicle is under UAV command. The pilot does not control the specific angle of bank for a turn or pitch and power adjustments for climbs and descents; they are controlled by an onboard computer system. The pilot acts ***in-the-loop*** in the sense that flight commands need to be entered. The UAV does not determine its own flight route. The UAV example illustrates how different system capabilities have the human in different roles depending on what actions are performed. In addition to the role differences, one can also appreciate how increasing the automation on the UAV moves the human further in the direction of an on-the-loop interaction.

When a UAV needs to operate beyond or without a continuous communication link, the humans that interact with the vehicle move further towards complete on-the-loop interaction with the UAV. Capabilities beyond flight control surface movements have to be automated. A human may send a command to the UAV to fly into an area, but once communication is severed, the UAV must be capable of determining the route to fly. Further, because the UAV has a mission beyond just flying, processes performed by sensor operators must also be automated. A vehicle that can determine and control its own flight path is of no use other than providing a distraction on an adversary's radar screen unless it can operate its own onboard sensors and eventually provide input to the overall battle space. If sensor interpretation is not included in the system's capabilities but instead has only sensor data storage, one or many human sensor operators will have to interpret the raw data into useful information once communication is reestablished. This can be cumbersome for the operators, and in situations where time is critical, it can delay operations that will use the information.

Advances in automation have shown the ability to improve processes and potentially reduce human workload. A growing field of research even aims to produce systems that can perform human "thinking" or cognitive functions. While these advances in computer programming may produce promising results through the execution of complex algorithms, humans still have to interact with the technology at some level. Because humans and "smart" machines are different, the interface space between the human and machine is a critical area of concern. There exists a need for humans in the interaction to accept the machine's processes as acceptable for task completion. If not, the machine will not be used as it was intended or possibly not used at all. Applying automation in human-machine systems requires the humans in the interaction to trust the automation. This thesis investigates the importance of trust in a human-machine system that includes processes that are carried out by automation substituting the processes of a human.

C. RESEARCH OBJECTIVE

The objective of this investigation is to develop trust need statements and system requirements that are traceable to functions in a process performed by a human-automation system. This is performed by leveraging and applying the findings of a thorough literature review to the functional process model of an "autonomous" aerial vehicle system conducting an intelligence, surveillance and reconnaissance (ISR) mission with the goal of positively identifying a vessel at sea. The model functions are categorized as either reasoning or non-reasoning functions performed by the system without communicating with a human. With the functions categorized, each were assessed for potential impacts on human trust. Finally, based on the trust evaluations, system capability needs and corresponding system requirements were developed.

D. THESIS ASSUMPTIONS

The following assumptions apply to the "autonomous" aerial system used in this study:

- Automation is fielded with a defined mission. There is no intent in this study to imply that the system creates its own mission. Automation is used to augment an existing manned system or tele-operated UAV.
- The limitations set forth in DOD Directive 3000.09 will not be violated. Specifically, the capabilities allowed by the directive will scope the process model that is used in this work. DOD Directive 3000.09 states that "systems that are onboard or integrated with unmanned platforms must be designed such that, in the event of degraded or lost communications, the system does not autonomously select and engage individual targets or specific target groups that have not been previously selected by an authorized human operator." (2012, 3)

E. THESIS ORGANIZATION

The following is a brief description of each chapter included in this work and provides a synopsis of the contents of each chapter as well as an overview of the thesis developments.

The current chapter introduces the area of investigation and identifies thesis objectives and assumptions. Chapter II is a literature review into the how the terms

automation, autonomy, autonomous, and automated are used to describe human-machine systems and how their use can affect capability descriptions and design. Chapter III investigates what trust is and how trust can affect the use of automation as well as identifying where trust in automation plays a role in system design. Chapter IV describes the process for identifying ships through use of a functional model and identifies the functions in the process where automation replaces human functions. Chapter V analyzes the process functions defined in Chapter IV for trust needs based on the findings in Chapter III and develops trust requirements for the modeled process. Chapter VI summarizes the findings and provides suggestions on how this work may influence future research.

II. AUTOMATION AND AUTONOMY DEFINITIONS

The use of the words “automated,” “automation,” “autonomy” and “autonomous” are sometimes used in place of each other to describe the same subject. Generally, each of the terms are used in reference to a system that includes a computer sub-system that can perform processes on its own with limited to no human direction. While replacing any of the terms with another may be acceptable in day-to-day conversation, the same is not true when considering writing requirements for military acquisition. Various publications recognize the importance of the use of the words and provide specific guidance on their meaning. However, when investigating published literature on the same subjects, the terms automated system or automation, and autonomous system or autonomy are used differently. This chapter will survey the various uses of the terms and provide a summary for their use within this thesis.

The literature survey begins with publicly released Department of Defense directives and reports.

A. DEPARTMENT OF DEFENSE PUBLICATIONS

One specific area in which the definitions for automated systems and autonomous systems are of clear concern is in their use when describing a weapons system. The Department of Defense Directive 3000.09, *Autonomy in Weapons Systems*, released in 2012 provides guidance in the following definition:

> Autonomous weapon system. A weapon system that, once activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapons systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation. (Department of Defense 2012, 13–14)

This definition’s key point is that autonomous weapons systems can select to engage targets; that is, the choice is made by the system not the human. The definition allows for systems that are monitored by a human who has the authority to deselect targets or stop the engagement process, but the human does not have a “vote” in the

initial decision to select an object for engagement. The directive defines human-supervised autonomous weapons systems as those "designed to provide human operators with the ability to intervene and terminated engagements" (Department of Defense 2012, 14). While specific to weapons systems, these definitions imply that a non-human system that can make decisions infers that system is an autonomous system.

The DOD directive gives further guidance in its definition of semi-autonomous weapons systems. They are defined as weapon systems "that, once activated, [are] intended to only engage individual targets or specific target groups that have been selected by a human operator" (Department of Defense 2012, 14). According to the directive, the sole function that makes a weapons system less than an autonomous one is that "human control is retained over the decision to select individual targets and specific targets for engagement" (Department of Defense 2012, 14).

The definitions provided by the DOD directive are intended to give guidance to the design and development of armed systems, but the clear distinction between autonomous and semi-autonomous relies on a specific decision capability. The views expressed by the directive indicate that there is a specific point in which systems are classified differently. While the terms automated and automation are not used to distinguish the differences, other DOD resources do use these terms for comparison.

Prior to the release of DOD Directive 3000.09, the Defense Science Board (DSB) released an information-gathering report identifying roles of and gaps in the current and future use of autonomous systems. The report begins by defining autonomy as "a capability (or set of capabilities) that enables a particular action of a system to be automatic or, within programmed boundaries, 'self-governing'" (Defense Science Board 2012, 1). The definition is purposeful in its use of the word capability. Within the study, "the task force reviewed many of the DOD-funded studies on "levels of autonomy" and concluded that they are not particularly helpful to the autonomy design process" (Defense Science Board 2012, 4). The argument that the DSB puts forth is that systems that include autonomy (a capability) should not be defined based on how much, or which functions are allocated to a computer. The report suggests that a level definition leads to an assumption that "there are discrete levels of intelligence for autonomous systems, and

that classes of vehicle systems can be designed to operate at a specific level for the entire mission" (Defense Science Board 2012, 4).

The 2012 DSB report advocates deserting levels of autonomy (or levels of automation as will be discussed later in this chapter) as metrics of measure for system design. The report acknowledges that "system autonomy is a continuum from complete human control of all decisions" (Defense Science Board 2012, 4) to a level in which much of the deciding is done by a computer-based system that may or may not be supervised by a human at a very high level. In the author's view, it is not that automation level descriptions are not valuable, but instead that using the levels as overall system descriptions is not. The intent is to avoid writing requirements that indicate a system shall be designed to a specific autonomy level. This can lead to designs that focus too much attention on the computer in the machine, instead of on the "collaborative" human-machine system and also implies that the computer can have a "discrete levels of intelligence" (Defense Science Board 2012, 4).

It is worth noting that the DSB uses the word "intelligence" in their argument for not using levels to describe autonomous systems. This gives the implication that autonomy over automation involves a system that determines or decides to perform an action; one that is able to perform cognitive or thinking (intelligence-based) functions. This is similar to DOD Directive 3000.09 on weapons systems in the view of decision making without human input. As with the DOD directive, the DSB report indicates that autonomy is different from automation.

Another Defense Department report specifically identifies a difference in automation and autonomy. In May 2015 the DOD Office of Research & Engineering's Autonomy Community of Interest released a report focusing on Test & Evaluation and Verification and Validation challenges for automated and/or autonomous systems. Their definitions are directly from an Air Force Research Laboratory strategy report (Masiello) and are:

> Automation: The system functions with no/little human operator involvement; however, the system performance is limited to the specific

> actions it has been designed to do. Typically, these are well-defined tasks that have predetermined responses (i.e., simple rule-based responses).
>
> Autonomy: The system has a set of intelligence-based capabilities that allows it to respond to situations that were not pre-programmed or anticipated (i.e., decision-based responses) prior to system development. Autonomous systems have a degree of self-government and self-directed behavior (with the human's proxy for decisions). (2013, 3)

These definitions, like the previous definitions, note that the ability to carry out intelligence-based or cognitive functions is what distinguishes autonomy. These specific definitions go further by implying that autonomous system decision making replaces human decision making. If the function is completed by using a model that represents a human's reasoning process, the system has autonomy. For the process to be defined under automation, no thinking or deciding is modeled. Instead the system just follows a series of steps to complete a task without deciding to complete it, or "think" or "reason" about how to complete it.

The most recently released DOD publication regarding automation and autonomy from the Defense Department was released in August 2016. The *Defense Science Board Summer Study on Autonomy* summary begins with the acknowledgement that "autonomy has many definitions and interpretations. For this reason, the report begins with an introductory section that defines the term and its context for the purposes of this study" (Defense Science Board 2016, 3). The given definition states that "autonomy results from delegation of a decision to an authorized entity to take action within specific boundaries. An important distinction is that systems governed by prescriptive rules that permit no deviations are *automated*, but they are not *autonomous*" (Defense Science Board 2016, 4).

Again the distinction between automated and autonomous lays in the ability of a system to make decisions. This report, however, implies that autonomy constitutes both automated and autonomous capabilities. The distinction continues with the DSB citing a presentation by L.G. Shattuck who uses a definition from the Institute for Human & Machine Cognition (IHMC): "to be autonomous, a system must have the capability to independently compose and select among different courses of action to accomplish goals

based on its knowledge and understanding of the world, itself, and the situation" (Defense Science Board 2016, 4). If the system has prescriptive rules related to the decision making then the system is not autonomous; if there are no such rules, the system is autonomous. However, the DSB continues, citing Bradshaw et al. (2013, 57) that "no machine–and no person–is truly autonomous in the strict sense of the word" (Defense Science Board 2016, 5) and that the report would use the word autonomous to describe capabilities and not systems.

Lastly, the Department of Defense Assistant Secretary of Defense Office for Research and Engineering in their 2015 technical assessment on autonomy considers the use of automation and autonomy effectively the same. In a footnote for their description of the broad fields of autonomy (DOD Research and Engineering) they state:

> While some in the field draw a distinction between automation as more rigid and autonomy as able to operate under higher level of complexity or uncertainty, we do not differentiate between the two because the distinction falls away when viewing them as enablers of DOD capabilities. Even if there is a difference of degree between them, they both perform the same function of enabling non-human decisions and actions. (2015, 1)

This view is primarily held because the amount of automation (or autonomy) can change for the same system depending on the functions in the process that are being completed and how much or where human interaction exists.

Recapping the Defense Department publications, the consensus is that automation and autonomy are different when describing systems. When describing high-level military capabilities, the view is that the two terms and their derivatives do not add value to defining a capability need.

At high levels of description, autonomy and automation may very well both just be enablers of a desired capability as the DOD Office of Research and Engineering states. However, the terms matter when the human-machine interactions are described and interfaces are designed. What a system can do and with how much authority must be understood by the humans involved in the interaction. Describing the functions as automated or autonomous to the human can give differing perceptions, especially when the function carried out is replacing a human cognitive process.

The DOD literature is not all that drives the understanding and development of systems that are automated or autonomous. The definitions shown thus far both have similarities and differences to descriptions offered by other sources discussed next.

B. NON-DOD PUBLICATIONS FROM HUMAN FACTORS, ERGONOMICS AND COGNITION FIELDS OF STUDY

The following examples investigate non-DOD specific literature on how automation and autonomy are defined.

According to T. Sheridan, one of the most widely cited authors on the subject of humans and automation, the "term's [automation] first use is traceable to a 1952 *Scientific American* article" (Sheridan 2002, 9). In his book *Humans and Automation*, he cites three definitions for automation from the Oxford English Dictionary and provides a breakdown of each of the three before expressing his own three-point "contemporary" definition:

> Automation refers to (a) the mechanization and integration of the sensing of environmental variables (by artificial sensors); (b) data processing and decision making (by computers); and (c) mechanical action (by motors or devices that apply forces on the environment) or "information action" by communication of processed information to people. (Sheridan 2002, 9)

This definition does not put boundaries on how little or how much "processing and decision making" must occur to consider a system automated. In fact, "processing" and "decision making" can be very different capabilities. Processing can imply the ability to sort data. Decision making can imply having the ability to choose from available (or collected) data. Processing can also imply having the ability to understand data. Likewise, decision making can imply comparing data. While different implications are clear, the definition is not complete without understanding Sheridan's view on how automation is characterized when the automation and human interact.

In 1978 Sheridan, along with Verplank, published a table titled: Levels of Automation in Man-Computer Decision-Making (Sheridan and Verplank 1978, 8–17) in their work on tele-operating an undersea vehicle. Using ten levels, the table describes how the human-automation interaction can vary. At level one, the human tells the

computer what and how to do everything. At level ten, the computer does everything, if it decides to, and tells the human that the process is complete, again if it (the computer) decides to tell the human. The intent of the table is "for stating what effect the inclusions of certain hardware or computer-based features" (Sheridan and Verplank 1978, 8–15) would have on the interaction space at the human-machine interface.

Sheridan's ten levels of automation are slightly modified in *Humans and Automation* to eight "degrees of automation" (Sheridan 2002, 62). The updated, simpler scale aims to join similar actions described in the ten levels and changes the wording for each of the degrees to emphasize the interaction between the human and the machine (computer). Both scales are models used to describe the range in which interactions change when different tasks are given to automated machines.

It is the author's view that the levels, or degrees, of automation developed by Sheridan were not intended to be used as "design to" levels but instead to be used as guidelines to describe the control level of the automated machine and the communication that occurs between the human and machine as automation processes are increased. In the DSB's 2012 *Task Force Report: The Role of Autonomy in DOD Systems*, the authors recommend that the DOD should "abandon the use of levels of autonomy" (Defense Science Board 2012, 4) because in their view the levels place too much attention on the machine design and not on the human-machine interaction (Defense Science Board 2012, 4). The report specifically makes mention of Sheridan's levels of automation in its assessment (DSB 2012, 23–24). Their wording agrees with the author's assertion that the levels should not be used as "design to" levels or as a categorical description for an entire system. However, the use of levels or degrees can be useful to help describe the desired roles the non-human and human actors in the system fill but, the numbering scheme with the descriptions is not useful. The numbers give the implication of static states along the automation continuum. Additionally, the use of the word automation instead of autonomy in Sheridan's work is notable because he does not establish a difference.

Sheridan is not alone in describing these systems with the word automation. Parasuraman and Riley, both also widely cited authors on the subject of humans and automation, use the term as well. Their definition follows a similar vein as Sheridan's:

> We define automation as the execution by a machine agent (usually a computer) of a function that was previously carried out by a human. What is considered automation will therefore change with time.…Today's automation could very well be tomorrow's machine. (Parasuraman and Riley 1997, 231)

From this definition, automation is anything that a machine does that replaces a human. This wide definition would include both of Masiello's definitions on automation and autonomy mentioned before. Using this definition also includes Sheridan's view on automation and the description of the different levels. Parasuraman and Riley, unlike Sheridan, use the word autonomy to help describe automation, in a model format.

Citing previous work by Riley, automation levels are described by the intersection of "intelligence" and autonomy (Parasuraman and Riley 1997, 232). The model (Riley 1989, 126) uses seven intelligence levels that range from "raw data" in which no processing is conducted to "operator predictive" in which operator actions and errors are anticipated. The autonomy scale has 12 levels ranging from "none" to "autonomous," indicating that the automation has varying levels of self-control. From Riley's model, automation is something that is designed. The automation is given a certain amount of "intelligence" or ability to perform actions, and it has a certain level of autonomy that describes the amount of control the automation has in performing the intelligent action.

Riley's model does not necessarily disagree with Masiello's definitions as they are written. The "intelligent" functions are performed by the automation while the capability to perform these functions with or without human interaction is described by autonomy. Conceivably, one could design a system with varying levels of automation and autonomy at different times. The functional processes (in many cases carried out by the execution of computer code) can be designed to any given level of automation, and depending on the desired capability, varying levels of autonomy can be afforded to the system to carry out the processes.

Continuing their investigations into levels and types of interaction between humans and machines and how function allocation can be described, Parasuraman and Sheridan joined with Wickens and wrote an article for IEEE in May 2000. Within the article a familiar, yet slightly modified, definition for automation is used and the use of

levels to describe interactions between humans and automation is again provided but also updated. Here, they define automation as "a device or system that accomplishes (partially or fully) a function that was previously, or conceivably could be, carried out (partially or fully) by a human operator" (Parasuraman et al. 2000, 287). The definition is very similar to Parasuraman and Riley's 1997 definition.

The levels provided in the IEEE article follow a similar style to Sheridan's previously published levels and are labeled as the Levels of "Automation of Decision and Action Selection" (2000, 287). The description of the levels here however, are described as "representing increased autonomy of the computer over the human action" (Parasuraman et al. 2000, 287) based on Sheridan and Verplank's original ten level scale. The terms automation and autonomy are used similarly to Riley's use for autonomy and Sheridan's use for automation. Instead of separating the two as different, autonomy is used to describe the role that the automation takes in the human-automation interaction. Parasuraman et al. continue by reiterating that automation levels do not remain constant for all processes the automation performs. They illustrate this by modeling a high-level functional process of information processing. In their four-step process, which is fashioned after a very simplified model of a human information processing method, they show how automation levels can vary across information acquisition, information analysis, decision selection and action implementation (Parasuraman et al. 2000, 288). The example model agrees with the DSB assertion that levels should not be used to describe systems as a whole. However, to properly describe the task distribution between computers and human in each process, the level descriptions are beneficial to highlighting what capabilities are designed into the system.

The Parasuraman et al. definition has been used by others. Most notably, and relevant to the research of this thesis is Madhavan and Weigmann's 2007 work that compares human-human trust to human-automation trust. Because of their topic, the definition used considers automation the full or partial replacement of functions that are carried out by a human (Madhavan and Weigmann 2007, 279). Another widely cited work by Lee and See concerning trust in automation, defines automation as "technology that actively selects data, transforms information, makes decisions, or controls processes"

(Lee and See 2004, 50). This definition infers self-governance much like other definitions of autonomy.

Depending on the source, automation and autonomy can be used to describe the same, or distinctly different types of systems. One common theme that stands out however, is that automation is typically used to describe system functions, while autonomy is used to describe system capability. Still, when considering the development of systems for military use, and specifically the writing of requirements, the use of these terms can matter. Their use also matters when developing training and operating literature for the people that will interact with these systems.

C. SUMMARY

The system, no matter how automatic, self-governing, or intelligent, will always at some level, interact with a human. The author agrees with the DSB and Bradshaw et al. that no system that is developed is ever fully autonomous. Through design, systems can perform a variety of functions spanning multiple levels of description with differing levels of authority to start, stop, continue or potentially to modify how or when the functions are performed. These are unique characteristics of these advanced systems.

This thesis will use the term automation to describe the functions an advanced non-human system performs. Autonomy will be used to describe how much oversight or direction is or is not exercised by the human on the advanced system. More autonomy refers to less oversight or direction and vice versa.

The system used in the analyses within this thesis is characterized as an "autonomous" aerial system. The intent of this thesis is to investigate where and how trust factors into the design of the automation in the "autonomous" aerial system so that the system will be correctly used and relied on. The next chapter investigates trust in automation.

III. TRUST IN AUTOMATION

This chapter addresses trust and its importance in automation. As systems include more advanced functionality the interactions between humans and the automation, as well as the allocation of the functions, in the system changes. Trust becomes a key need in these systems because of the changes. Parasuraman (1997, 236) affirms, "trust often determines automation usage." Similarly, Sheridan (2002, 77) states that "trust is now considered an important factor in human-automation performance." The changes in the functional roles and the interactions cause differing perceptions and expectations of the automation by the human. The need for trust in human-machine military systems is identified as a major challenge by the Department of Defense Research & Engineering's Autonomy Community of Interest on Test and Evaluation, Verification and Validation (ATEVV) working group in their 2015 Technology Investment Strategy Report. Identified as current challenge 4, the ATEVV writes:

> Handoff, communication, and interplay between operator and autonomy become a critical component to the trust and effectiveness of an autonomous system. … When [modeling and simulation] is not possible at design time, how can trust in the system be ensured, what factors need to be addressed, and how can transparency and human-machine system requirements for autonomy be defined? (Department of Defense 2015b, 4)

To determine what requirements are needed to build trust into these systems, an investigation into what trust is and what can influence a human's trust in automation is needed.

A. DEFINING TRUST IN AUTOMATION

A large body of work exists concerning trust in automation. Much of the published work investigates the factors that affect trust in human-human relationships and applies, or attempts to apply, the findings to human-machine relationships. Lee and See, paraphrasing others, state, "Sheridan (1975) and Sheridan and Hennessy (1984) argued that just as trust mediates relationships between people, it may also mediate the relationship between people and automation" (Lee and See 2004, 51). This cross agent assertion that factors affecting trust in human-human relationships can be used to

determine factors affecting trust in human-automation relationships is useful. However, it should *not* be stated that trust in the human-human relationship is the same as trust in the human-automation relationship. Humans and machines are different. While they are different, one actor in the relationship is a human; therefore, understanding factors that affect trust in the human-human relationship are beneficial to determining factors in the human-automation relationship.

Before determining the factors that affect trust, the roles of the human and the automation in the relationship must be assigned. A relationship in the simplest form requires two actors, and when considering trust, the actors can be labeled the Trustor and the Trustee. For the simplest human-automation relationship the Trustor is the human and the Trustee is the autonomous system. Generally, trust in this relationship is directional; the Trustor exercises trust (or lack of trust) in the Trustee.

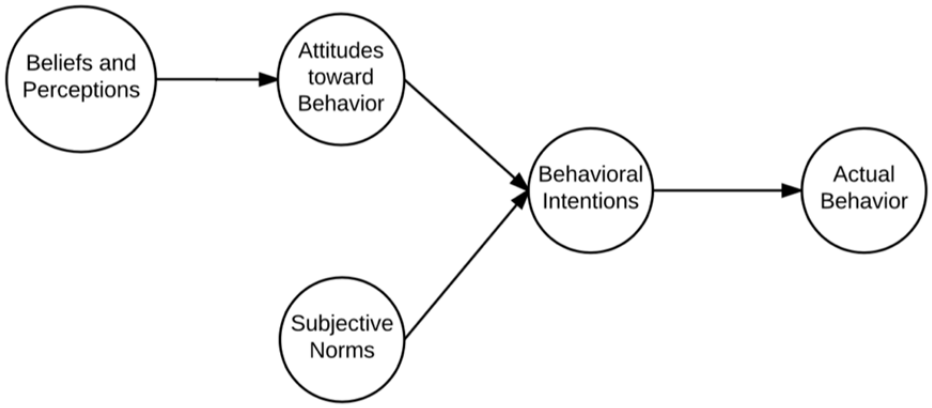
To describe one way in which a Trustor exercises trust in the Trustee, Lee and See (2004) summarize the conclusions of several other researchers that focus on trust as an attitude or expectation of the Trustor in the Trustee. Each example identifies that the Trustor has a perception (preconceived or initial) of the Trustee. The Trustor expects and depends on a certain response or action from the Trustee. Lee and See identify through comparison of the examples that "trust concerns an expectancy or an attitude regarding the likelihood of favorable responses" (Lee and See 2004, 53). Madhavan and Weigmann (2007, 280), like Lee and See, also encapsulate some of the more widely cited sources on trust. Their summary states: "Trust refers to the expectation of, or confidence in, another and is based on the probability that one party attaches to co-operative or favourbale [sic] behaviour [sic] by other parties" (Barber 1983, Muir 1987, Hwang and Buergers 1997). From these examples it is implied that trust in automation is an attitude that the human (Trustor) has toward the automation (Trustee) based on the what the trustor expects or believes the automation will do in a given situation. However, an additional approach in describing trust is common that does not describe trust as an attitude.

Instead of describing trust as an attitude of the Trustor based on the Trustee's actions, trust has also been explained as an intention or willingness to act in uncertainty. Again summarizing some widely cited definitions on trust, Lee and See identify that

many authors identify “vulnerability” as a critical element of trust. For trust to be an important part of a relationship, individuals must willingly put themselves at risk or in vulnerable positions by delegating responsibility for actions to another party” (2004, 53). The vulnerability approach to trust describes trust as an intended behavior of the Trustor.

Further still, Lee and See identify that “some authors go beyond intention and define trust as a behavioral result of vulnerability or risk (Deutsh, 1960; Meyer, 2001)” (2004, 53). This definition explains trust as an outcome produced by the vulnerability of the Trustor. Therefore, depending on which group of authors one reads, trust is either a belief, an attitude, an intention, or a behavior. Lee and See perfectly point out that “these distinctions are of great theoretical importance, as multiple factors mediate the process of translating beliefs and attitudes into behaviors” (Lee and See 2004, 53). The Defense Science Board Summer Study on Autonomy, quoting Hoffman et al. (2013, 84), also correctly acknowledges that “trust is complex and multidimensional” (Defense Science Board 2016, 14) Investigating the literature shows that this is clearly true.

Thankfully, Lee and See’s investigation into trust does not stop at identifying the inconsistencies that are found in the body of literature. They continue by using Ajzen and Fishbein’s Theory of Reasoned Action to help resolve differences in the identified definitions of trust. In the Theory of Reasoned Action, behaviors are described as the results of a person’s intentions, and their intentions are effects of attitudes. A person’s attitudes are influenced and formed by their beliefs, perceptions and knowledge (Lee and See 2004, 53). Figure 1 shows a simplified model of these associations.



Behaviors are the result of intentions, which are functions of attitudes that are influenced by beliefs, perceptions and knowledge.

Figure 1. A Generalized Model of Ajzen and Fishbein's Theory of Reasoned Action. Adapted from Ajzen (2007).

Using the Theory of Reasoned Action allows for addressing the inconsistencies observed in the definitions used for trust by investigating reliance in a system and how reliance and trust are related. Lee and See assert that "trust stands between beliefs about the characteristics of the automation and the intention to rely on the automation" (Lee and See 2004, 54). Agreeing with this assertion, the author provides the following statement that adds on to that used by Lee and See; it is used as the definition for trust in the remainder of this thesis.

> Trust is the attitude of a human, developed from beliefs, perceptions and knowledge of a system's functional capabilities, towards the behavior of reliance in the system's actions to achieve the human defined goals in situations characterized by uncertainty and vulnerability.

From the definition, trust influences behavior; and, human behavior is one of the many considerations Human Factors Engineering (HFE) practitioners investigate when designing for human use (Lockett and Powers 2003, 463). In Chapter II, it was stated that no system is ever fully autonomous; a human will interact with and use the automation. Therefore, HFE considerations, and thus trust, are important to the system design. The following section discusses a few of the more widely cited investigations into trust in automation.

B. INVESTIGATIONS INTO TRUST IN AUTOMATION

The behavior of reliance on automation has been investigated. Parasuraman (1997, 236) cites two examples in which reliance, or perceived reliability, is related to a person's trust through examples of empirical studies. In the first example, Parasuraman cites a study performed by Muir (1988) on the use of an automated aid to control a simulated beverage manufacturing plant (1997, 236). He discusses how Muir argues that if a system is perceived as honest and trustworthy, then a person's trust is positively affected, and the person will rely on the system. Similarly, if the person experiences let down or betrayal (a belief that the system is not as reliable as originally thought), then the person's trust is negatively affected, and the system is relied on less. Muir also argues that if trust has been negatively affected, it takes time to reestablish (Parasuraman 1997, 236).

The second example referenced by Parasuraman (1997, 237) and also by Madhavan and Weigmann (2007, 284) summarizes the findings of Lee and Morray (1992). Lee and Morray (1992) performed an experiment that investigated the relationship between trust and control strategies for humans interacting with a semi-automatic pasteurization plant. The analysis on their experiment, like Muir's (1988), showed a correlation between subjective trust and reliance on automation. In a subsequent study, Lee and Moray (1994) later determined that conditions in which a person has the ability to choose to perform a task or to allow an automated machine perform the same task, reliance on the automation is higher when trust is higher (Parasuraman 1997, 237). The findings in Lee and Morray's (1994) study also indicate that a person's self-confidence in their own ability to complete a task affects the reliance that is placed on a machine completing the same task (Parasuraman 1997, 237). They found that trust in the automated machine is higher when the perception of one's own ability to complete the task was lower than the perception of the automated system's ability. Additionally, the opposite was found to be true; if one's own self-confidence in task completion is high, then the automation would not be used.

According to the research performed by Muir (1988) and Lee and Moray (1992), if the beliefs and perceptions a human has towards an automated system's performance

and reliability are affected negatively, trust in automation is slow to recover if trust recovers at all. On the other hand, Parasuraman (1997, 237) provides an example study by Riley (1994) in which the slow recovery of trust does not occur. Describing Riley's findings, he explains: the operators "did not delay turning on automation after a recovery from a failure [of the system]; in fact, many participants continued to rely on the automation during the failure." Additionally, Parasuraman in his own investigative study with others (Parasuraman et al. 1994) found similar results; that participants continue to rely on the system's automation even after the system failed.

Parasuraman's examples do not necessarily contradict the example shown by Lee and See. Instead they both highlight the importance of beliefs, perceptions and knowledge and how they affect trust and reliance on automation. The studies show that the attitude of trust is scalable. Based on perceptions, beliefs and knowledge, conditions of distrust and over trust can be reached resulting in the rejection of automation or over-reliance in automation respectively. Parasuraman and Riley's 1997 article titled "Human and Automation: Use, Misuse, Disuse, Abuse," provides multiple examples of these undesirable cases.

These extremes imply that the attitude of trust must be calibrated to the system's capabilities for a human to rely on it appropriately. Parasuraman and Riley state, "If automation is to be used appropriately, potential biases and influences on this decision [to use or not use automation] should be recognized by training personnel, developers, and managers" (1997, 238).

C. CALIBRATING TRUST IN AUTOMATION

Lee and See (2004, 55) provide a description that explains the connection between trust and automation capability (trustworthiness). They use the term resolution to describe "how precisely a judgment of trust differentiates levels of automation capability (Cohen et al. 1999)." The goal in using automation is to align the human's trust with the capability of the system, or achieve calibrated trust where "trust matches system capabilities, leading to appropriate use" (Lee and See 2004, 55). What their description does not directly show is that the human's trust level is influenced by the perception,

belief, or knowledge of the system's "trustworthiness." This is an important factor. While trust calibration can be defined by the appropriate intersection of trust level range and system capability range, one must be careful not to assume that a high level of capability results in a high level of trust or visa-versa. Instead, the description highlights a goal or result space. When considering a system with a certain set of capabilities, designers must aim to influence the human's trust level to a calibrated level by influencing the beliefs, perceptions and knowledge of the system's capabilities.

The system capabilities must be clearly explained to, and understood by, the human that will interact with it in order to influence the human's beliefs, perceptions and knowledge of the system. The methods in which actions are performed, and how the system will respond to errors must also be understood. Finally, the human in the interaction must have a reason to expect that the autonomous system will perform its actions to meet the intended objective. The methods used to address these areas can vary. The humans that interact with an automated system can have diverse levels of knowledge, skills, and abilities. Humans will also have varying levels of experience interacting with differing levels of automated functionality in systems and may be predisposed initially to reject system process that differ from the status quo for a non-automated process to complete the same action. Others may, on the other hand, be predisposed to accept the autonomous system blindly because of an impression that the system must be better than the status quo or else it would not be presented for use. Neither of these attitudes is desirable, especially in cases where automation is used in military operations, lifesaving operations, or other high-risk applications.

It is imperative that actions are taken to calibrate trust in automation appropriately. As stated before, trust in human-system relationships is often derived from examples of trust in human-human relationships. However, "people are often thrust into relationships with automation in a way that forces trust to develop in a way different from that in many relationships between people" (Lee and See 2004, 67).

Lee and See, summarized by Madhavan and Weigmann (2007, 280) identify that the performance, process and purpose, of the automation are bases that influence trust. That is, these three bases can influence the knowledge, beliefs and perceptions of the

human. Performance, both past and present, includes the automation's reliability, predictability and ability. Process refers to the algorithms that the automation uses and how well they explain how the automation functions. Purpose refers to why the automation is designed.

D. DESIGNING TO CALIBRATE TRUST IN AUTOMATION

One sure way to affect the level of trust that people have when they interact with automation is through training. It has already been established that trust is influenced by the knowledge that a person has about the automated system. It should be expected that if a human is to interact with an automated system, that education and training are available, and often required, before any interaction with the automated system occurs. Education and training can give the human a baseline knowledge of the system's performance, process and purpose. While important, education and training are not the only methods to affect a person's trust attitude. Lee and See (2004, 67) agree: They state, "Thus, early in the relationship, trust can depend on purpose and not performance. The specific evolution depends on the type of *information provided by the human-computer interface* and the documentation and training" (emphasis added).

The design of the automated system must allow for the calibrated trust attitude to remain calibrated. As mentioned before, it has been shown that a change in the human's perception in the automated system based on the interaction experience can have undesirable effects on the trust attitude. This means that the knowledge gained from education and training, acquired before interacting with automation, is not by itself sufficient to keep trust calibrated and thus appropriately affect reliance. Lee and See state, "Like trust between people, trust in automation develops according to the information available, rather than following a fixed series of stages" (2004, 67). The interfaces that exist between the human and the automated machine are where information is made available which makes them critical points where trust must be addressed.

The presentation of the material in an ergonomics sense is important. Lee and See address this when summarizing the findings on two studies into trust and credibility in

computing as follows: they state, “In many cases trust and credibility depend on surface features of the interface that have no obvious link to the true capabilities of the system (Briggs, Burford, and Dracup 1998; Tseng and Fogg, 1999)” (2004, 73). However, the interface must also take into consideration what the information is and how it was developed before determining how it is presented. Not all information equally carries the same trust needs. The best designed interfaces allow for information exchange in such a way that it aides in the proper calibration of trust.

Depending on the capabilities of the automation, different types of information are exchanged at the human-automation interface. Information that is the result of working out equations is easily traceable and understood. Education and training often are sufficient to allow for information based on these calculations to be trusted when the results are presented. For example, one is taught in classical physics that speed (velocity) is the change of position over a given time. Even without calculus, one can be taught that the measured distance between two points divided by the elapsed time to move from the first point to the second will provide the speed of the moving object. This type of information is automatically calculated and displayed in a number of systems that humans interact with in everyday life. Such interfaces include speedometers used in personal vehicles and airspeed indicators as well as altimeters used in aircraft cockpits. Depending on the importance of the presented value to the human, the location and method of reporting matters but displaying or communicating the method of calculation is not necessary at the interface. Information that is solely based on equations with input variables obtained from the measurements of position, distance, time, pressure or temperature for example, only require the result of the equation to be reported. This, however, does not alleviate the need for error reporting. If the sensors used to collect the variable data, or the computer used to calculate the resultant information are damaged, or produce errors, the human must be informed at the interface. Unless the equation is chosen, or created by the computer, the knowledge of how the equation is calculated prior to interaction is sufficient to accept the reported value as true.

Automation’s sole use is not to perform calculations and report results. Automation includes the employment of algorithms to replace thinking or reasoning type

actions that are normally performed by humans. The algorithms that are developed may provide very accurate results when compared to humans performing the same action; however, they are not exact replicas of the human's process. They are instead abstractions created by a programmer. Some cognitive functions that are carried out by a machine's algorithms may use the very same information that is expressed through the results of the aforementioned equations. Combining equation-based information through the use of algorithms is carried out through coded programs based on any number of matching or decision-making models and requires more than the final result to influence trust. The equation-based information may be clearly understood and accepted as accurate but, how and why the information is put together in an algorithm may not be fully understood by the human. Lack of understanding can lead to rejection, blind trust, or may cause confusion.

In a human-human relationship, the Trustor may have an opportunity to ask the Trustee how a conclusion was ascertained and also may sense how much trust should be placed in the Trustee's information based on the Trustee's voice or body language (perception of the Trustee). In human-automation relationships the interaction does not allow for the same observations unless some sort of anthropomorphic features or process explanation are designed into the machine to emulate these observables. The algorithmic generated information then requires more than the final value or result in the exchange.

The interface design must communicate information to the human in such a way that the human can determine if the presented information matches known and perceived machine capabilities and the situation. Lyons (2013) identifies the same necessity in one of his three models for human-robot transparency factors. He states, "the analytical model [for transparency] needs to communicate the underlying analytical principles used by the robot to make decisions" (Lynn 2013, 50). The human may not be able to query the automation how or why a presented solution was given in a traditional sense, and displaying the actual algorithmic steps requires knowledge of the chosen coding language and time that may not be afforded in many cases. Because of these limitations, it is imperative that the human-automation interfaces communicate algorithmic generated data in a descriptive manner that fits the operational relationship between the agents. For the

interface to communicate in this way, the automation processes that use such algorithms must be identified, and requirements must be written for the additional needed descriptive information beyond the algorithm result.

In many uses of automation, the systems are not intended to be viewed as human-like. Instead, these systems are viewed as advanced tools that humans use in a team setting. The system becomes an integral part of the team but is not necessarily a team member in the same respects that a human is a team member. In an investigation into the challenges of including the advanced systems that have been the topic of this chapter as team members, Klein et al. (2004) eloquently state that the development of these systems exists at a boundary between two schools of thought. The first is that these systems are developed with the goal of creating systems "that emulate human capabilities" (2004, 94). The second being "to create systems that extend human capabilities, enabling people to reach into contexts that matter for human purposes" (2004, 94). In the author's view, this boundary or "fine line" as described by Klein et al., can begin to be bridged with calibrating trust at the human-machine interface focusing on the information exchange that includes algorithmic generated data.

E. SUMMARY

In regards to trust in the human-automated system relationship, trust is an attitude that the human has toward relying on the automated system in an environment of vulnerability for use in reaching a goal. Trust must be calibrated such that humans will used automated systems appropriately within the system's designed capabilities. The trust attitude exhibited by the human is variable. It is influenced by, and calibrated through the knowledge, perceptions and beliefs that the human has toward the automated system. The proper calibration of trust is executed by both training and through system design; neither alone can maintain the appropriate level of calibrated trust. Education can establish a baseline for the human's trust attitude in the automation. Training and system design can aide in maintaining properly calibrated trust. System design affects trust at the human-automation interfaces. The exchange of information at the interface and how the information is presented is key in addressing the human's perceptions and beliefs about

the automation and hence influences trust. According to the Theory of Reasoned Action, the trust attitude will then affect the behavior of reliance. The type of information exchange at the interface carries differing trust needs. Information based on sensor data output and provable calculation alone, which are presented as values, influence the trust attitude through perception at the interface in an ergonomics sense and through knowledge of the equations used. Information based on the modeling of human decision making, cognitive processes, or matching algorithms influence the trust attitude through perception of their outputted results and also through knowledge and perception of the methods by which the results were developed. A need to communicate the methods along with the results exists. Therefore, the actions performed by automation in a system need to be defined and then scrutinized for how they can affect the human trust attitude. One method for doing this is through a functional analysis of the system (Lockett and Powers, 2003, 469).

The next chapter elaborates on the scenario introduced in Chapter I and uses a functional architecture to decompose and define the automation actions an “autonomous” aerial vehicle must perform it to complete the scenario mission. Chapter V then analyzes the actions for trust needs discussed in this chapter.

IV. IDENTIFYING AUTOMATION TRUST POINTS IN A FUNCTIONAL PROCESS

A. INTRODUCTION

The first chapter of this thesis introduced a capability gap in using a UAV to extend the operational awareness of a ship at sea. The gap exists because the UAV relies on a communication link for the human operators to deliver flight guidance commands and also for tele-operating the onboard sensors. In order to bridge this capability gap with a ship-launched unmanned aerial system, it is necessary to include automation in the system. Operating beyond the communication link requires giving the system some amount of autonomy (within operational and design constraints) to control both flight functions and sensor operating functions. As described in Chapter II, all systems, regardless the "automation level" or how much autonomy the system has, at some point will interact with a human. A system designed to fill the capability gap, while given the authority to perform functions on its own, could never be fully autonomous. The system design and operating instructions are created by humans and are limited to the allowances of the programmed algorithms.

Recognizing the necessity for increased automation in a system to fill the capability gap introduces issues of relying on the proposed system and accepting its ability to replace human operators. In Chapter III, reliance was described as a behavior that is influenced by trust, an attitude, through the application of the Theory of Reasoned Action. It was found that trust is influenced through the knowledge, beliefs, and perceptions the human has of the automation. According to Lee and See (2004), and summarized by Madhavan and Weigmann (2007, 280), the automation's design can affect the knowledge, beliefs, and perceptions through the three bases: purpose, performance and process.

For the proposed "autonomous" aerial vehicle, the purpose, or why the automation is designed (Madhavan and Weigmann 207, 208), is to conduct an Intelligence Surveillance and Reconnaissance (ISR) mission in place of a manned helicopter beyond the limits of a communication link. The human(s) interacting with the

automation must understand the automation's purpose, and how the system functions will meet the purpose. Education and training prior to interaction with the system can affect the knowledge of the human and develop a trust foundation. This trust baseline will affect how the human uses the system. The purpose of the system must be known before use to expect proper use and proper beliefs in the system's capabilities.

The performance of the automation, defined by the ability of the automation to meet the purpose, and can be described by reliability and predictability (Madhavan and Weigmann 207, 208). Both knowledge of the reliability determined through test and evaluation of the system and the perceived reliability experienced through interaction with the automation will affect trust. Predictability of the automation's functional sequence affects trust through perception. When, why, and in what order processes are completed must follow a sequence that is understood by the human interacting with the system.

The process of the automation refers to the algorithms used to complete the functions. As described in Chapter III, the algorithms relate to knowledge, beliefs, and perceptions through more than the final result determined by the algorithm. Lee and See (2004) suggest that algorithms used in automation should be either simple, or revealed in such a way that they are comprehensible. A suggested method is to show intermediate results as the algorithm runs (2004, 74). For the proposed aerial system, revealing intermediate steps is not possible without a communication link. Furthermore, because the automation's functions include those that are modeled after human cognitive processes, the algorithms used in the system would certainly be very complex.

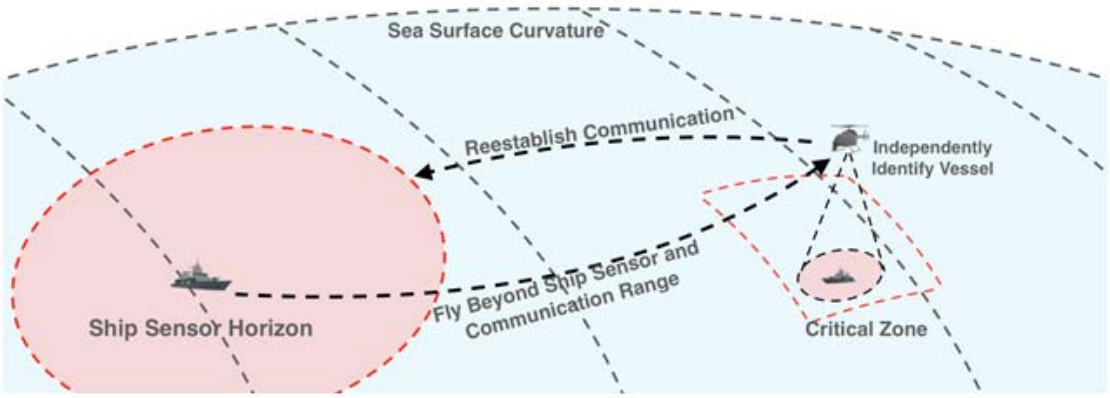
While complex, and inherently lacking the ability to have process transparency, explaining and communicating automation processes to users must still be addressed in the system design to allow proper calibration of user trust and reliance in the proposed system. This requires an investigation into the automation functions used to meet the desired capability. With the functions identified, points for simplification or for explanation can be identified, and design requirements that are specific to calibrating trust in the automation can be generated. While the human-automation interface is the point at which perception of the automation's performance and process is affected, the

interface cannot do so without receiving the necessary information from the system. Where the information that affects trust is generated must be determined.

The remainder of this chapter utilizes a functional architecture, modeled with Functional Flow Block Diagrams (FFBDs), to decompose the ISR mission. The FFBDs used are based on the functional architectures developed by Frau, et al. in "An Architecture for an Autonomous, Weaponized Unmanned Aerial System (UAS)" and by Hernandez, et al. in "MH-60 Seahawk / MQ-8 Fire Scout Interoperability" as well as the author's own experience with manned aerial ISR missions.

B. MISSION DESCRIPTION

The mission for an aerial system was briefly introduced in Chapter I with the identification of the capability gap. A visual description of the desired mission capability is provided in Figure 2.



An aerial vehicle is launched from a ship and flies beyond the Ship Sensor Horizon and communication ranges to independently search for and identify a vessel. After identifying a vessel, the air vehicle moves to reestablish communication and transmit the findings to the ship. This diagram is for informational purposes only and is not drawn to scale.

Figure 2. Mission Description Diagram

The operational awareness of the crew on an air-capable ship sailing on the open ocean is limited to the sensor horizon of the ship's onboard sensors and the available information that may exist in a tactical network from other ships or aircraft. To increase awareness beyond the sensor horizon, an aerial vehicle is launched on an ISR mission

with the specific purpose of locating and identifying other vessels. The aerial vehicle is unmanned and is capable of performing the ISR mission without constant communication with the ship. When a vessel is located and identified, the vehicle will reestablish communication and send a contact report to the ship. When mission time expires, or fuel load on the vehicle requires returning to the ship, the vehicle flies to the ship and lands.

Without communication with the ship, the vehicle must utilize automation to navigate and maneuver, by interpreting the environment and controlling the vehicle on its own. The methods used by the automation to perform these capabilities undoubtedly require human understanding and calibrated trust for reliance in the system. However, even if these processes are adequately automated, understood, and trusted, automation must also be incorporated to operate vehicle sensors and also for interpreting the sensor data. The sensor data could arguably be stored in the vehicle's computers and transmitted back to human sensor operators or data interpreters on the ship after the vehicle reestablishes a communication link. Yet, it would better resemble the current capability provided by manned helicopters if the vehicle could send back interpreted information. Automation in sensor employment and data interpretation are critical to meet the capability gap. Therefore, the process architecture will focus on the sensor and data interpretation functions the vehicle must perform.

In Chapter I it was mentioned that a helicopter such as the MH-60R would typically be used for this type of ISR mission. The automation used to replace the helicopter then must, at a minimum, have the same sensor capabilities as the helicopter. No ship-launched UAV currently exists with the same sensor capabilities as the MH-60R. The purpose of this study is not to develop such a system. However, it would not be a stretch of the imagination to visualize such a vehicle. The current inventory of military UAVs includes systems that employ visual and infrared (IR) cameras, radars, and electro-magnetic signature surveillance sensors. Therefore, for the purposes of this work, an assumption is made that the vehicle's sensor suite includes a surface search radar, visual and IR capable cameras, an Electronic Support Measure (ESM) system, a data link, and a GPS system; all of which are systems installed on the MH-60R and are used in an ISR

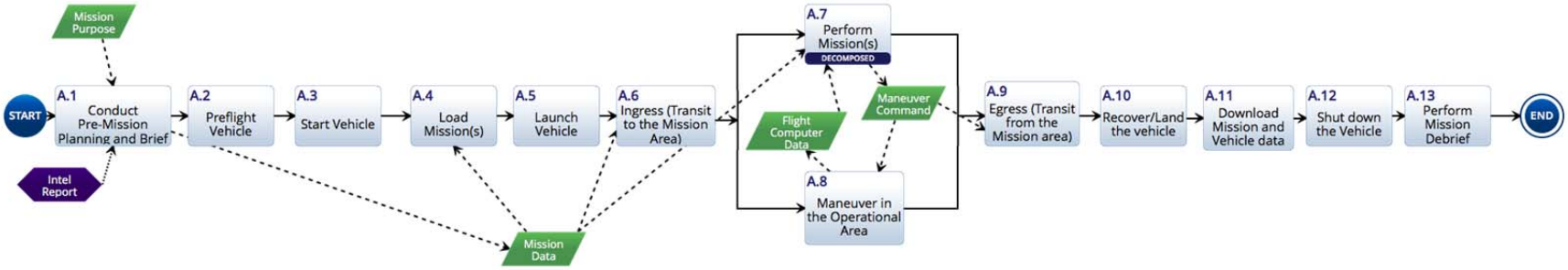
mission. The vehicle also includes the necessary computers systems to operate the sensor suite.

The described vehicle is not a typical UAV. Some of the human-machine interactions with the vehicle may at times be similar to those a system operator would have with a UAV; but, when operating without the communication link, the system is closer to the DOD Directive 3000.09 definition (US Department of Defense 2012, 14) of a semi-autonomous system. The literature review provided in Chapter II discussed the importance of properly and clearly using the terms autonomous and automation, so a description is warranted for this system that considers the terms. The system is an aerial vehicle that includes automation in its design and the algorithms used by the automation include those that represent human cognitive processes. The system is an Aerial Vehicle with Automated Cognitive Capabilities (AVACC). To distinguish the vehicle from existing UAVs and to avoid unintended perceptions that rise from the use of the term autonomous, the acronym AVACC will be used to describe the system and as the system name.

The functional processes of the AVACC ISR mission to locate and identify vessels on the sea without a communication link are shown and described in the following sections of this chapter. Each decomposition level of the model is introduced in separate sections and the section headings include the model decomposition number scheme. The section numbers and letters do not follow the model numbering scheme.

C. PERFORM ISR MISSION

A model describing the high-level functions for the AVACC ISR mission is shown in Figure 3. The model is adapted from models produced by Frau et al. (2011), and Hernandez et al. (2010) as well as the author's own experience operating military aircraft.



High level functional description of Aerial Intelligence, Surveillance and Reconnaissance Mission.

Figure 3. Perform ISR Mission High Level Functions

The entire functional process is not decomposed in this investigation. The purpose of the investigation is to identify functions in the process where human-automation trust needs exist. As mentioned before, the focus in this investigation is on the high-level function Perform Mission (A.7). The mission is expected to occur in an environment where communication with an authoritative boundary system (the ship) is not present. At the highest level, the model can be used to describe a large number of aerial vehicle missions. While the referenced projects influence the model, many of the similarities are due to the commonality in describing aerial vehicle operations.

To perform any mission, a few prerequisite resources are needed. For the modeled process, a Mission Purpose and Intelligence Report must be provided to the first function, Conduct Pre-Mission Planning and Brief. The output from the Mission Brief function is Mission Data. The Mission Data is a required resource for the functions Load Mission, Ingress (Transit to Mission Area), and Perform Mission. One constraint for this overall process is that a mission has to exist. The AVACC does not search for or create missions. Much like a manned aircraft system, a purpose must be defined and sufficient data must be collected before establishing the need for a mission. For this investigation, an assumption is made that a mission exists and that the necessary information for a mission brief has been provided. It is possible that automated technology may aid in the mission development process, but that capability is beyond the scope of this investigation.

In addition to the prerequisite items, the Perform Mission function (A.7) interacts with the function Maneuver in the Operational Area (A.8) through the resources Flight Computer Data and Maneuver Command. Because the system is intended to operate without communication, the AVACC must be capable of maneuvering for the mission and to reestablish communication with the ship. The Perform Mission function (A.7) does not include the functions required to maneuver the AVACC, but does create maneuvering needs based on observations, calculations and vessel identification. This is modeled with the output Maneuver Command from the Perform Mission function to the Maneuver in the Operational Area. Additionally, Flight Computer Data is used in Perform Mission sub functions. This data is created under the Maneuver in the Operational Area function

(A.8). The Maneuver in the Operational Area function (A.8) is not decomposed in this investigation.

D. OVERVIEW OF PERFORM MISSION(S) (FUNCTION A.7)

The Perform Mission function (A.7) of the process is based on the find, fix, track, target, engage, and assess (F2T2EA) kill chain that is a commonly used description for military system missions and is shown in Figure 4. Per DOD directive 3000.09 (2012, 3), the AVACC cannot be used to select and engage targets on its own; and, the Perform Mission function (A.7) is completed beyond the communication range of the ship. Therefore, functions after and including Target are not decomposed in the model.

Perform Mission is adapted from the common F2T2EA kill chain model.

Figure 4. Perform Mission (A.7)

The F2T2EA process has been slightly modified. Find is replaced with Search Area, Fix is replaced with Detect, and the function Resolve Contacts is added. Search Area replaces Find because finding an object is the desired result of searching. The action that the AVACC performs is searching and the model intent is to show the actions (functions) that describe the mission process. Detect replaces Fix in the model because detecting an object, like Find, is also a result. A fix describes where in tactical space an object was detected. Find and Fix describe the intended results of functions where Search and Detect describe what is being done. Track is retained in this process because it is the action that the AVACC performs to follow the position and movement of a detected object.

An object can be tracked without knowing what the object is and the object's movement can help with refining a classification, but classification and identification are

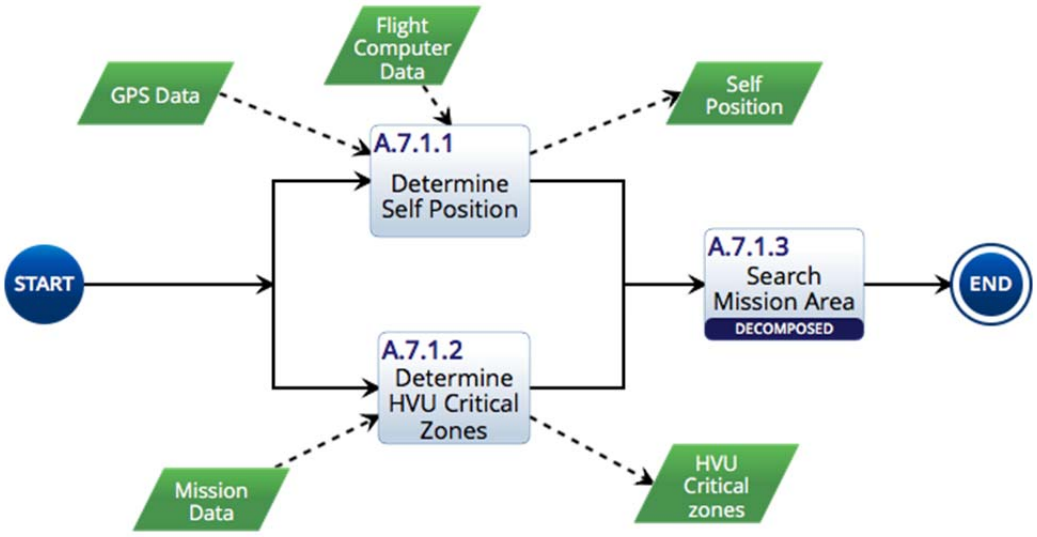
separate actions. The function Resolve Contacts is added to describe how the actions of the tracked object is classified and identified.

Even without the DOD Instruction 3000.09 (2012) prohibition for using the AVACC to select and engage targets, the resolution of contacts must occur and be trusted before continuing to the next step in the kill chain whether a human is involved or not.

The next sections in this chapter decompose functions A.7.1 through A.7.4 in greater detail.

E. SEARCH AREA (FUNCTION A.7.1)

The AVACC must determine and verify its own position and also must determine where the High Value Unit (HVU) Critical Zones are while searching the mission area. The HVU Critical Zones are defined by the Intelligence Brief and are included in the Mission Data. They are specific sub areas in the greater Mission Area. If a vessel is found in an HVU Critical Zone, the HVU may need to change its operating scheme depending on the found vessel's identification. In this scenario the only HVU is the "home" ship. How or why the ship would change its operating scheme is beyond the scope of this analysis. Multiple inputs and outputs are identified that influence the decomposition of Search Area and are shown in Figure 5.



Search Area (A.7.1) replaces Find in the common F2T2EA kill chain model.

Figure 5. Search Area (A.7.1)

For the function Determine Self Position (A.7.1.1), the system's position is determined with GPS Data as well as determining the time and distance from a previous known point; known as dead reckoning. With GPS or dead reckoning, the AVACC can use position, velocity and acceleration equations to determine where it is. Determine Self Position is a continuous function that is also included in the Navigate function, a sub function of Maneuver in the Operational Area (A.8) which is not decomposed in this investigation. The Flight Computer Data input comes from the Navigate function.

Determining the location of the HVU Critical Zone (A.7.1.2) is important to the search process because the HVU Critical Zones refine and prioritize the AVACC's search. Doctrine and Rules of Engagement should dictate how close any object should be allowed to come to the ship. This distance would be contained in the Mission Data created and loaded before the AVACC is launched and would establish the HVU Critical Zones. Once these determinations are made, the AVACC can search the Mission Area.

1. Decomposed Search Mission Area (Function A.7.1.3)

For the AVACC to gather information, an area of interest, described and refined by the HVU Critical Zones and Mission Data, must be defined and an optimal search pattern must be determined before data collection is performed by the onboard sensors. Figure 6 shows this.

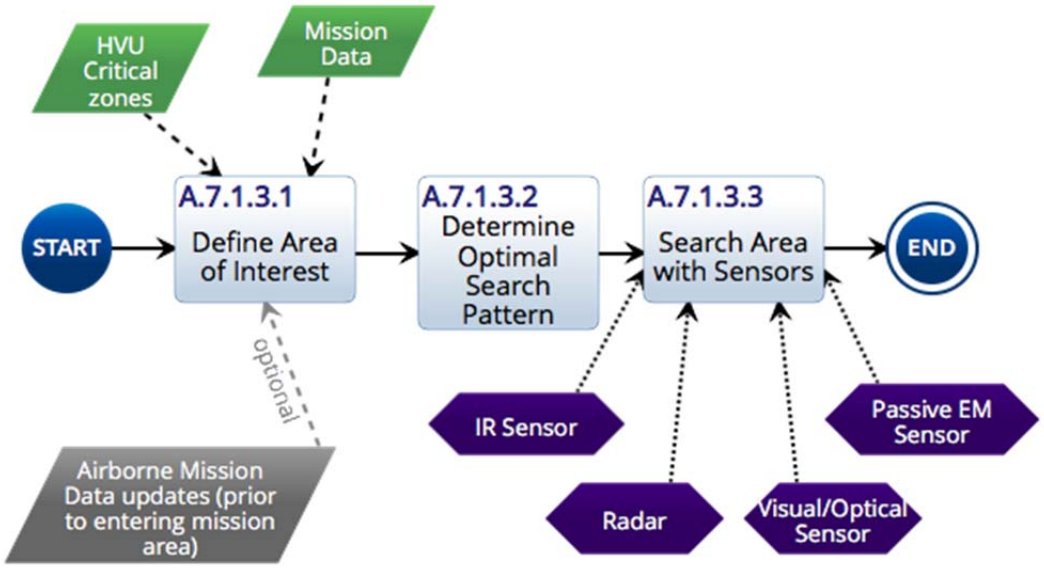


Figure 6. Search Mission Area (A.7.1.3)

From the loaded Mission Data and the determination of the HVU Critical Zones, areas that need to be searched are determined (A.7.1.3.1). Priority would be determined in the mission load. In situ uploads could also be used as inputs to defining the search area(s) of interest if received before communication with the ship is lost. The receipt of the Airborne Mission Data updates would have to occur during the AVACC Ingress (Transit to the Mission Area) (A.6) while communication is still established. The resource is modeled as optional in this process because it is not required for the mission. It is possible for mission changes to occur after the AVACC is launched and the optional resource is modeled to show this. The purpose of showing the three inputs to the Define

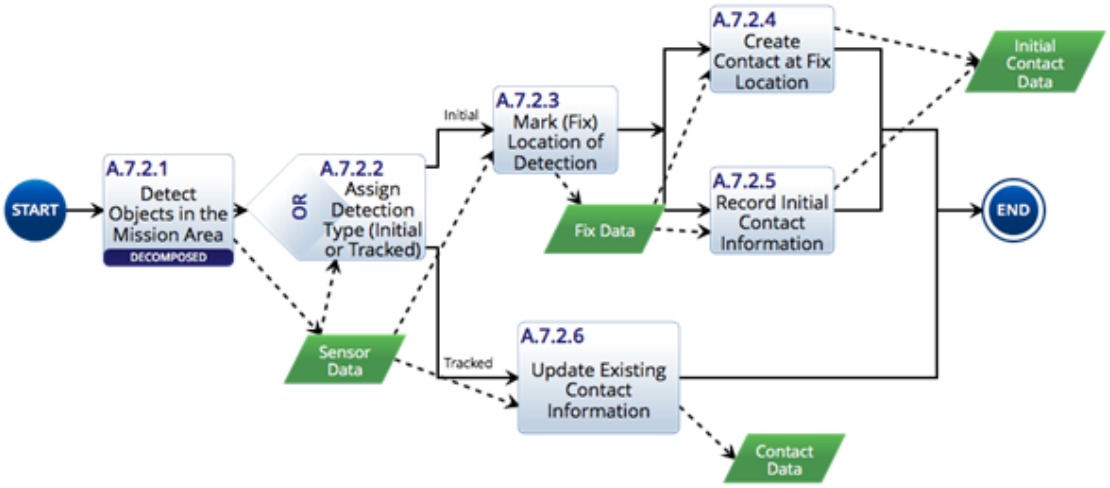
Area of Interest function is to show that the AVACC is operating on an assigned mission. It does not have the capability of creating a mission on its own.

The function Determine Optimal Search Pattern (A.7.1.3.2) is completed similarly to the method used to determine search and rescue search patterns. The Mission Data includes a starting point based on the HVU Critical Zones. The pattern is generated by running an optimization algorithm to cover the HVU Critical Zones. The same automated process is leveraged in manned aerial systems.

The system sensors include an IR sensor, a radar, a visual sensor, and a passive EM sensor to actively and passively search the area (A.7.1.3.3) with priority given to the Critical Zones.

F. DETECT (FUNCTION A.7.2)

Objects in the mission area that meet sensor threshold criteria for positive detection are marked for further investigation. The overall Detect process is shown in Figure 7. The Detect function is a prerequisite to and also a sub function of the later Track (A.7.3) function.



Detect replaces Fix in the common F2T2EA kill chain Model

Figure 7. Detect (A.7.2)

Detect Objects in the Mission Area is decomposed further to describe what a detection is and how it is completed by the AVACC. The decomposition includes the sensors involved and the resource links to and from other functions. This is shown in the next section with Figure 8.

1. Decomposed Detect Objects in the Mission Area (Function A.7.2.1)

Each of the AVACC's installed sensors have predetermined, designed thresholds when met, determine that a detection has occurred. These thresholds are set by the system capability requirements which are based on the intended use of the system. Once sensors are installed very limited adjustments can be made to the thresholds. Modes of sensor operation, for example radar scan rate, or visual and IR sensor optical zoom, can change a detection threshold. The detection criteria are determined by the mode and type of sensor in use. This describes the Define Detection Criteria function (A.7.2.1.1) in the decomposition of Detect Objects in the Mission Area shown in Figure 8.

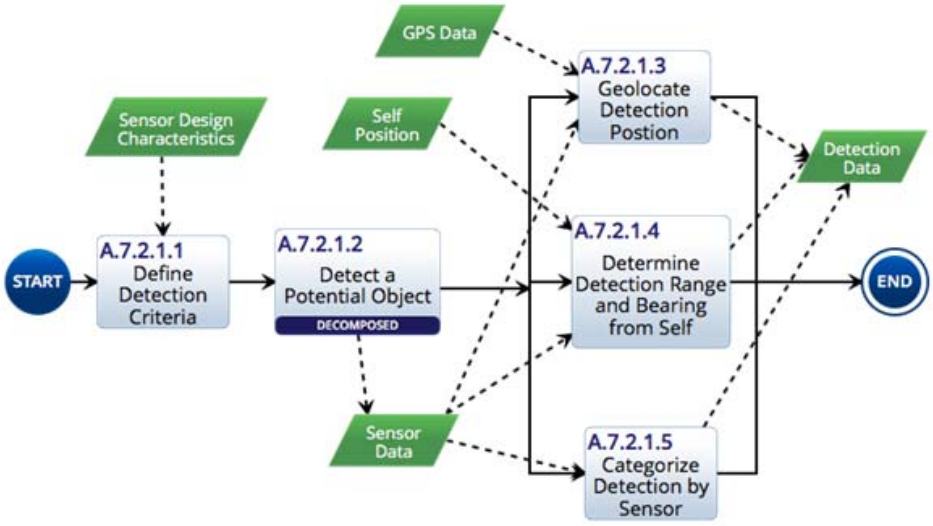


Figure 8. Detect Objects in the Mission Area (A.7.2.1)

Detect a Potential Object, function A.7.2.1.2, is decomposed further to show the sensors, sensor actions (functions), and information output from the sensor that defines a detection in Figure 9.

a. Detect a Potential Object (Function A.7.2.1.2)

The functions describing Detect a Potential Object (A.7.2.1.2) are shown in Figure 9. The sensor(s) that observe the detection create Detection Data which is used to create the initial Fix.

Typically, radar, and to a lesser extent ESM, are used for initial detection. Optical and IR sensors are typically used to investigate existing detections when range to the detection is reduced. It is possible to use Optical sensors and IR sensors for initial detection, especially when observing for events such as fires, but in the case of searching for vessels at sea they are not the primary search sensor.

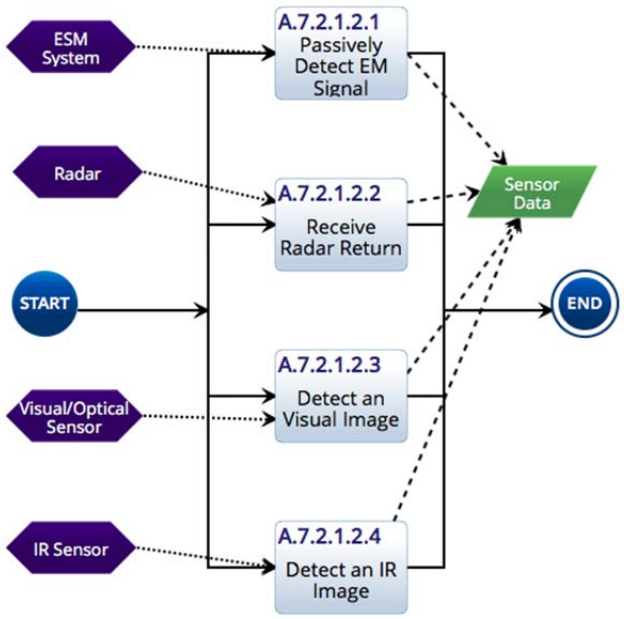


Figure 9. Detect a Potential Object (A.7.2.1.2)

The selected sensors are typical for a manned ISR mission. In order to replace a manned aircraft system, the AVACC should have at a minimum, the same sensors installed to perform the ISR mission.

Detections from the ESM system (A.7.2.1.2.1) are triggered by receiving an electromagnetic (EM) signal that matches a signal of interest listed in a library file included in the mission load. This is no different than how ESM is used on manned aircraft.

Radar generated detections (A.7.2.1.2.2) are no different than radar detections with a human operating the radar system. The radar transmits electromagnetic EM energy into space, the EM energy reflects off an object and returns to the radar receiver. If the reflected EM signal is above a designed threshold, a detection has occurred.

Detections form the Visual/Optical (A.7.2.1.2.3) and the IR sensors (A.7.2.1.2.4) on the AVACC are different from those in manned systems. In the manned system, a human is alerted to the possibility of an object in view by observing the shapes and colors in the sensor's display. In the AVACC, the sensor's view would be scanned electronically for contrasting shapes or colors that meet a threshold defining a possible detection. A series of algorithms would be used to perform this function. Google™'s image recognition software in their Cloud Vision API is one example of how automation can utilize algorithms to interpret visual information. Google™'s methods use the Internet as a source to help with image interpretation. The AVACC would not be connected to the Internet, nor would it have a communication link, so it would need an onboard database for its algorithms to use for matching and comparing.

b. Functions A.7.2.1.3–A.7.2.1.5

Geo-locating the Detection Position (A.7.2.1.3), Determine Detection Range and Bearing from Self (A.7.2.1.4), and Categorize Detection by Sensor (A.7.2.1.5) occur in parallel after an object is detected. Geolocation describes where the detection is in tactical space using Sensor Data and GPS Data. The bearing and range are determined using the Self Position and Sensor Data. Lastly, the Sensor Data is used to categorize the detection as a Radar, EM, Visual, or IR detection. These three functions together comprise the Detection Data.

2. Functions A.7.2.2–A.7.2.6

Following the Detect Objects in the Mission Area (A.7.2.1), the system assigns a type to the detection (A.7.2.2). This is modeled as an "OR" function as seen in Figure 7. The detection is either an initial detection or is associated with a contact that is already being tracked by the system. This distinction needs to be made because Detect (A.7.2) is also a sub function of Track (A.7.3).

The first time an object is detected, its position is recorded in a database as a Fix. In a manned system this can be automated but is typically performed by a human operator interacting with an interface. The operator places a symbol on the tactical display at the detection location. The Fix records the position and time that the detection occurred. Without a human in the system a detection would automatically be "marked" with a fix (A.7.2.3). It may not necessarily be a physical mark on a structured plot. Instead, a data point is recorded to a database with detection latitude and longitude, the time of detection, and the sensor associated with the detection.

A fix does not move. Therefore, the next step is in the processes is to create an entity that can move (A.7.2.4): a Contact. The Fix data is recorded (A.7.2.5) as the first data point for the contact.

The outputted resource of the Detect process is the Initial Contact Data. This data is updated (A.7.2.6) with subsequent detections in the Track function (A.7.3). Detection is required before tracking, but once a detection is made, the detection process becomes a step within the Track function.

G. TRACK (FUNCTION A.7.3)

Once the area has been searched and an object detected, a decision is made to track this specific object. The Track function is depicted in Figure 10. To distinguish the contact from other objects detected in the mission area, a unique label is assigned to the contact (A.7.3.1) and used to describe the contact in the database.

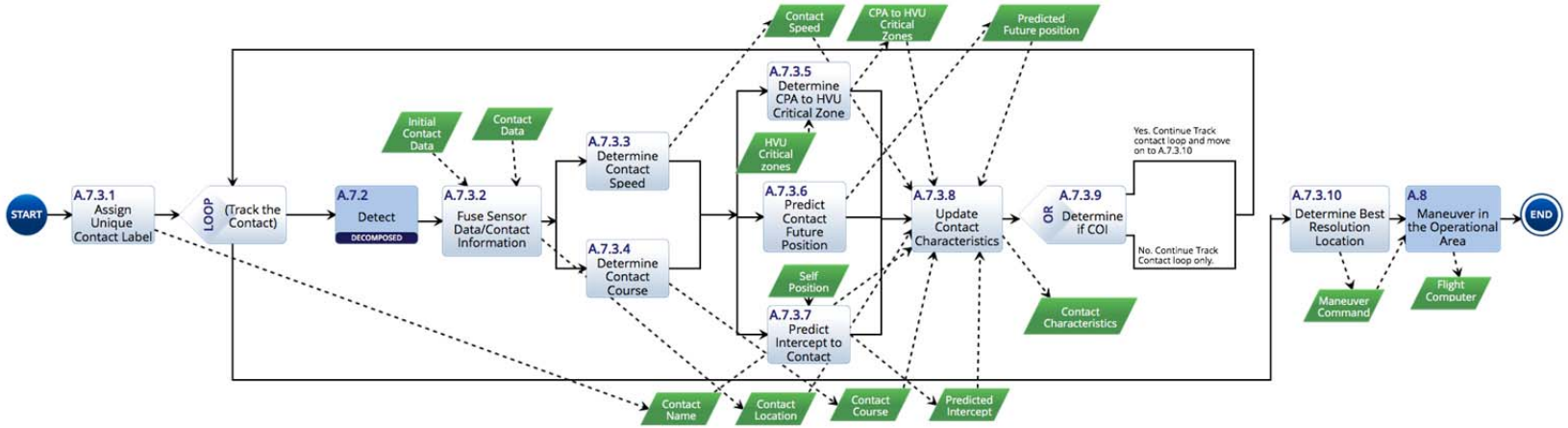


Figure 10. Track (A.7.3)

To track the labeled contact, subsequent detections from the sensors are fused and used to develop the measures of speed and course (A.7.3.2–A.7.3.4). With course, speed, and location determined a predicted closest point of approach (CPA) to the HVU Critical Zone, a predicted future contact position, and a predicted intercept point can be determined (A.7.3.5–A.7.3.7). Each of these characteristics are used to update the contact characteristics file (A.7.3.8). Based on the characteristics and known positions or zones of interest, the contact can be assigned another label: a Contact of Interest (COI) (A.7.3.9). A COI is a contact that is positioned or is moving in such a manner that it could potentially affect ship or HVU operations, or cause operational or defense concerns for the ship. The specific criteria for COI labeling would be included in the loaded Mission Data. Giving a contact a COI label increases the contact investigation priority over other tracked contacts. Tracking continues whether a contact meets COI criteria or not. This is depicted as a loop in Figure 10.

When a contact is labeled as a COI, the AVACC determines where it must position itself to resolve the COI's type, classification and identity (A.7.3.10) while continuing the track loop. This determination generates a maneuver command. The AVACC positions itself based on the best position determination. Maneuver in the Operational Area (A.8) is described as a flight control function that occurs concurrently with the entire Perform Mission (A.7) function as shown in Figure 3.

H. RESOLVE CONTACTS (FUNCTION A.7.4)

The next step in the standard F2T2EA process would be to target the tracked contact. However, targeting requires knowing more about a contact than its position, its speed and, where it is moving. Additional information is needed to describe what the contact is. Visual information is required to determine the contact type, its classification, and its identity. In manned systems, the visual sensor data interpretation is completed by a human operator. To fill the capability gap, the typing, classification, and identification is performed by the AVACC. Resolve Contacts (A.7.4), which is based on the processes a human operator would complete, is added in the modified F2T2EA process and is shown in Figure 11.

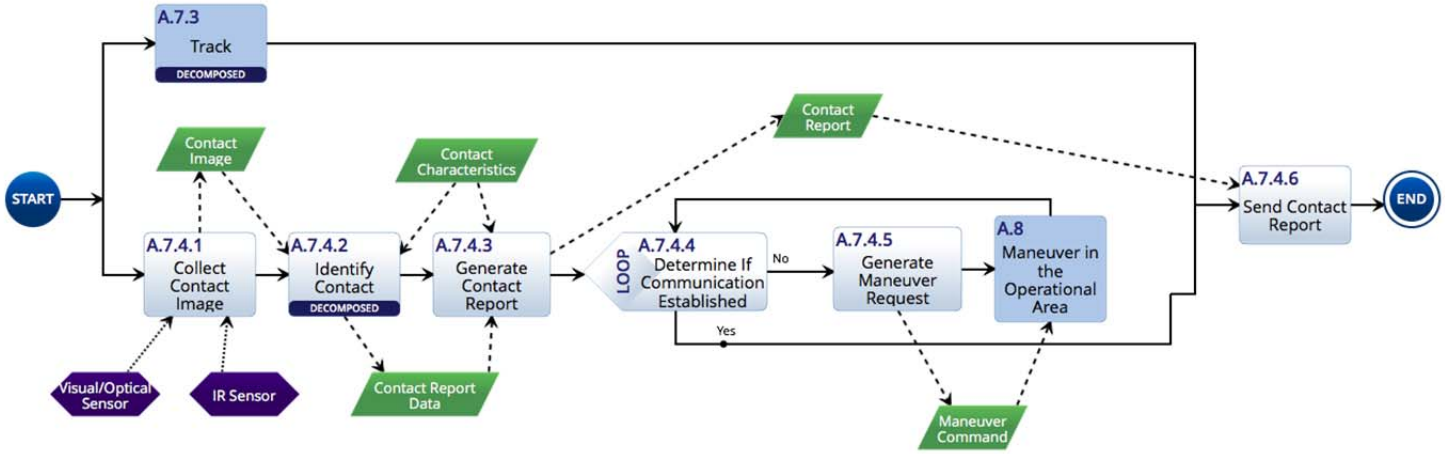


Figure 11. Resolve Contacts (A.7.4)

Contacts are further scrutinized in the Resolve Contacts process while tracking the contact continues. Resolving the contact primarily requires interpreting visual and/or infrared sensor data (A.7.4.1). Once an image of the contact is obtained it becomes part of the Contact Characteristics which the system uses to Identify the Contact. Identify Contact (A.7.4.2) is decomposed into sub functions and is shown in Figure 12.

1. Decomposed Identify Contact (Function A.7.4.2)

The functions that decompose Identify Contact (A.7.4.2) all include "decision making," matching, and comparisons type functions and are depicted in Figure 12. When performed by a human operator, the human's perception of the data, their knowledge of how the data is put together and their prior experiences with similar tasks are relied on to reach a final result. For the AVACC, the functions must be completed by executing algorithms that are model representations of the human operator's cognitive process. There are different ways to model a human's cognitive process; it is possible for completely different algorithms to perform the same task and achieve the same result. Because of this, information pertaining to the algorithm process must be included with the algorithm results in the human-AVACC interaction.

In the Compare Collected Data to Known Information function (A.7.4.2.1) the AVACC must determine what data in the Contact Characteristics is pertinent to ultimately identify the contact. The contact image is not used in this comparison. Algorithms compare collected data in the Contact Characteristics to pre-programmed mission briefs, intelligence reports and a pre-loaded Ship Recognition Knowledge Database that includes all potential vessels that may or may not be operating in the assigned mission area to develop a refined Possible Contact List. Once the collected data has been compared to known data the results of the comparison, the Possible Contact List, and the Contact Image are used to determine the contact type (A.7.4.2.2), classification (A.7.4.3), and identification (A.7.4.4).

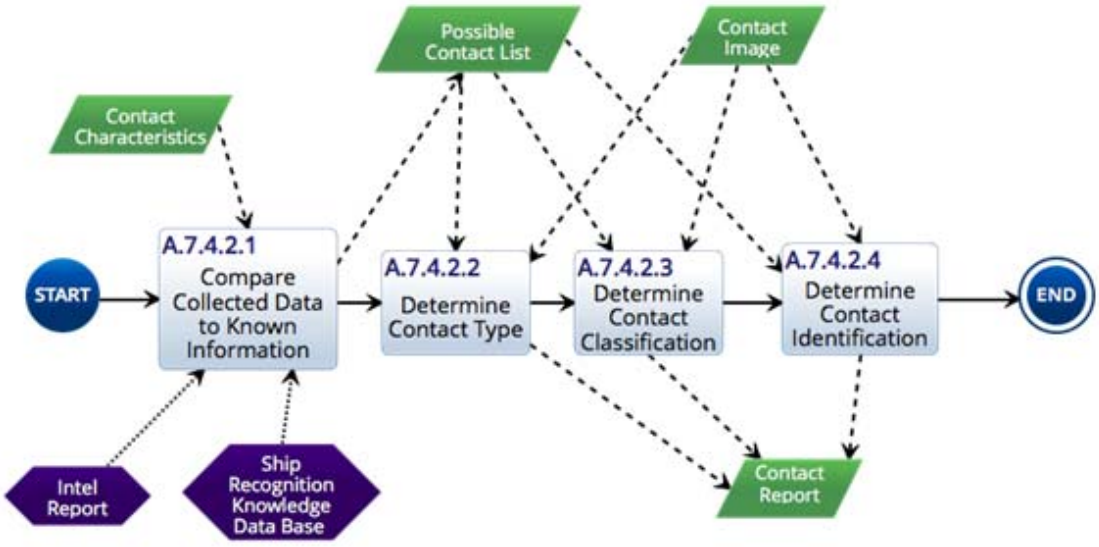


Figure 12. Identify Contact (A.7.4.2)

Determine Contact type (A.7.4.2.2) is performed by comparing the contact's size and speed to predetermined thresholds and the Possible Contact List. Based on speed and size, the contact type is labeled as a surface vessel, an aircraft, or potentially a subsurface contact's periscope. The AVACC would need to utilize image processing algorithms that examine the Contact Image for this function.

Classification of a contact (A.7.4.2.3) is resolved after the type is determined and is a further refinement of the contact description. Classification examples for contacts of the surface vessel type include warship, container ship, and fishing vessel to name a few. In a manned system a human operator would perform the classification by comparing the visual information to a known data base or experience with the imagery of different vessels. The AVACC would utilize image processing algorithms to execute this function.

Identifying the contact (A.7.4.2.4) includes assigning a country of origin, hull number, name or other detailed descriptions that identify the contact as unique. In a manned system this is performed by a human who compares sensor data, primarily visual, to a learned or available database of different vessels. The AVACC would again need image processing algorithms and a database of vessel images to execute this function.

Because determining the type, classification and, identification of the contact all involve a human cognitive processes, each of these three functions require more than the algorithms' generated results in the contact report. The algorithms need to be developed so that they provide results that are understood by the humans expected to use the information provided. Additionally, because intermediate algorithm steps cannot be observed without a communication link, justification for the results, explanation in how the results were determined, and the actual results must be included in the communication exchange via the Contact Report.

2. Functions A.7.4.3–A.7.4.6

The output from Identify Contact is a collection of information labeled Contact Report Data which used in the generation of the Contact Report (A.7.4.3). Once the Contact Report is generated, it is sent to the ship. It is important to understand that the Contact Report in this model is NOT the human-system interface. The Contact Report is a collection of information that is provided to the human-system interface.

A communication link must be verified or established before sending the contact report (A.7.4.4). If the AVACC is not in communication with the ship, a Maneuver Command is generated (A.7.4.5) and sent to the flight computer to move the AVACC into communication range (A.8). When communication is established, the contact report is transmitted to the ship (A.7.4.6).

The entire mission process occurs without human interaction or direct supervision. The Contact Report is the first, and possibly only, AVACC to human information exchange about the operational area beyond the ship's sensor horizon. When the contact report is transmitted, the AVACC has left the mission area. The ship cannot gain any further information about the mission area other than what is included in the Contact Report.

I. ISR MISSION FUNCTIONS NOT MODELED

As stated in sections C and D of this chapter, some of the higher level functions are not decomposed in this functional architecture. The following sections provide general descriptions of these functions.

1. Kill Chain Functions (A.7.5–A.7.7)

The Kill Chain functions after Resolve Contacts (A.7.4) do not fit into an ISR mission. Additionally, the restrictions stated in DOD directive 3000.09 (2012, 3) for "autonomous" systems prohibit targeting and engagement of contacts without human interaction. While the AVACC performs the first half of the modified F2T2EA Kill Chain (A.7.1–A.7.4), another system would be required to Target (A.7.5), Engage (A.7.6), and Asses (A.7.7) an identified vessel.

2. High Level Functions A.9–A.13

As mentioned in Section C of this chapter, the entire ISR mission process is not modeled; the focus is on the Perform Mission (A.7) function. With the Contact Report sent (A.7.4.6) to the ship, and communication reestablished, the AVACC would interact with the humans on the ship to complete the ISR mission by performing the remaining functions (A.9– A.13) which are depicted in Figure 3.

J. SUMMARY

The model provided in this chapter has explained the processes the AVACC must complete to meet the ISR mission. The model gives a small, specific example of what automation in a system would have to do if it replaced human sensor operator. Factors such as communication security, AVACC physical movement, and others are not modeled. The purpose for keeping the model simplified is to keep the focus on the trust needs related to AVACC identifying a vessel at sea. Further decomposition could change the focus to how an algorithm should be created to perform the actions. In many cases different algorithms could conceivably be written to complete the same function.

From the decomposed model it is clear that automation in the AVACC's design is critical to closing the capability gap. The AVACC's automation controlled processes are what must be trusted by the humans that assign the AVACC the ISR mission and also by any human that interacts with the AVACC or its Contact Report.

In the next chapter, the decomposed functions describing the Perform Mission function (A.7) are analyzed for how they affect trust in the AVACC. This analysis will identify the trust needs that can be used to write system hardware or software requirements for the AVACC. The decomposed Perform Mission functions are provided in Table 1.

Table 1. Perform Mission Functions Identified from Process Model

Number	Function Name
A.7.1.1	Determine Self Position
A.7.1.2	Determine HVU Critical Zones
A.7.1.3.1	Define Area of Interest
A.7.1.3.2	Determine Optimal Search Pattern
A.7.1.3.3	Search Area with Sensors
A.7.2.1.1	Define Detection Criteria
A.7.2.1.2.1	Passively Detect EM Signal
A.7.2.1.2.2	Receive Radar Return
A.7.2.1.2.3	Detect a Visual Image
A.7.2.1.2.4	Detect an IR Image
A.7.2.1.3	Geo-locate Detection Position
A.7.2.1.4	Determine Detection Range & Bearing from Self
A.7.2.1.5	Categorize Detection by Sensor
A.7.2.2	Assign Detection Type (Initial or Tracked)
A.7.2.3	Mark (Fix) Location of Detection
A.7.2.4	Create Contact at Fix Location
A.7.2.5	Record Initial Contact Information
A.7.2.6	Update Existing Contact Information
A.7.3.1	Assign Unique Contact Label

Number	Function Name
A.7.3.2	Fuse Sensor Data/Contact Information
A.7.3.3	Determine Contact Speed
A.7.3.4	Determine Contact Course
A.7.3.5	Determine CPA to HVU Critical Zone
A.7.3.6	Predict Contact Future Position
A.7.3.7	Predict Intercept to Contact
A.7.3.8	Update Contact Characteristics
A.7.3.9	Determine if COI
A.7.3.10	Determine Best Resolution Location
A.7.4.1	Collect Contact Image
A.7.4.2.1	Compare Collected Data to Known Information
A.7.4.2.2	Determine Contact Type
A.7.4.2.3	Determine Contact Classification
A.7.4.2.4	Determine Contact Identification
A.7.4.3	Generate Contact Report
A.7.4.4	Determine if Communication Established
A.7.4.5	Generate/Send Maneuver Request
A.7.4.6	Transmit Contact Report

Ultimately, the factors that affect, and can be used to calibrate, trust will be leveraged in the AVACC-human interaction at an interface. However, before the requirements for an interface can be developed, the sources of trust affecting factors must be identified and requirements for the AVACC must be written that ensure the factors are made available to the interface.

THIS PAGE INTENTIONALLY LEFT BLANK

V. ANALYZING PERFORM MISSION FUNCTIONS FOR TRUST NEEDS

A. INTRODUCTION

In Chapter III, Section D, a distinction was made between automation processes that are equation-based and those that are algorithmic-based. It was stated that the methods can affect trust differently. In the AVACC system, information is created, and actions are performed using both methods. However, while Chapter IV mentions using equations and algorithms for some of the modeled functions, the author acknowledges that the term "algorithm" is too broad when discussing the methods used by a computer to complete actions; a computer program is essentially a series of algorithms. Therefore, to properly address the differences between the functions, different descriptions are warranted.

Some of the functions that the AVACC must perform are normally completed by human thinking or decision-making processes in a manned system. Instead of using "equation-based" and "algorithmic-based" to describe the Perform Mission functions, they can be categorized by whether they represent human cognitive processes or not. The DSB (2016) discussion about the barriers to trust in automation provides an explanation to why these types of functions are unique. The report states:

> For some specific algorithm choices—such as neuromorphic pattern recognition for image processing, optimization algorithms for decision-making, deep neural networks for learning, and so on—the "reasoning" employed by the machine may take on a strikingly different path than that of a human decision-maker. (2016, 14)

The Perform Mission functions are then separated into two categories described by functions that are not "reasoning" functions (Defense Science Board 2016, 14) and those that are. This chapter will analyze all of the Perform Mission functions based on the Chapter III discussion on trust. How each of the Reasoning and Non-Reasoning functions affect human knowledge, beliefs and perceptions through the AVACC purpose, performance and process (Madhavan and Weigmann 2007, 280) will be addressed. The Non-Reasoning functions will be analyzed first followed by the Reasoning functions.

B. NON-REASONING FUNCTIONS

Of the 37 Perform Mission functions identified in the Chapter IV AVACC ISR mission model, 25 are Non-Reasoning functions. These 25 Non-Reasoning functions are further characterized as Database Read and Write functions, Equation functions, and Sensor functions.

1. Database Read and Write Functions

Ten of the Non-Reasoning functions are categorized as Database Read and Write functions. These functions either write sensor, equation, or Reasoning function data to computer files, or look up (read) data in the loaded Mission Data. The Database Read and Write functions are listed in Table 2. The Generate Contact Report function (A.7.4.3) was determined to be a special case of the Non-Reasoning Database Read and Write function category, and is analyzed separately in the next section.

Table 2. Non-Reasoning Database Read and Write Functions

Database Read and Write Functions		
Number	Function Name	Write or Read
A.7.1.2	Determine HVU Critical Zones	Read
A.7.1.3.1	Define Area of Interest	Read
A.7.2.1.5	Categorize Detection by Sensor	Write
A.7.2.2	Mark (Fix)Location of Detection	Write
A.7.2.3	Create Contact at Fix Location	Write
A.7.2.4	Record Initial Contact Information	Write
A.7.2.5	Update Existing Contact Information	Write
A.7.3.1	Assign Unique Contact Label	Write
A.7.3.8	Update Contact Characteristics	Write
A.7.4.3	***Generate Contact Report***	***Read/Write***

The Database Read and Write functions are part of the process that the AVACC performs to eventually identify a vessel, which is the purpose of the system. All of these functions describe actions that create data that will be used in the Reasoning functions, or are inputs to the Contact Characteristics. Those that are inputs to Contact Characteristics are included in the Contact Report as shown throughout the model in Chapter IV. They

do not have trust needs associated with how they are written or read; they basically describe computer bookkeeping. But, their inclusion in the Contact Report is necessary for fully describing a contact. Identifying a contact but not having information that describes its location, speed, and direction of travel, is not very useful.

Two functions are labeled as Read in Table 2. These functions can affect trust in the AVACC processes. For example, the Search Mission Area function (A.7.1.3), described in Chapter IV, Section F., includes an optional resource for the Determine Area of Interest (A.7.1.3.1) function. The optional resource, could conflict with, or be significantly different than loaded Mission Data. Therefore, the source of the read data should be communicated in the Contact Report. Reporting the source of read data provides knowledge of why the AVACC performed its search in a certain location. If a Contact Report was received that described a contact in an area that was not described by the loaded Mission Data, the perception of the AVACC processes are influenced negatively, and the human may be inclined to stop using the AVACC. This follows Parasuraman and Riley's (1997, 244–246) discussion of automation disuse. If the source of the read data that caused the AVACC to search in an area is included in the Contact Report, knowledge and perception of the AVACC's process is influenced which supports trust calibration. It may be appropriate to reject some AVACC generated data because computers do error. So, providing data to the human that can clarify unexpected discontinuities can increase the likelihood that the human will understand and accept the results. For the Database Read and Write functions, the easiest way to facilitate this is to define Mission Data by a Mission Number and include the Mission Number in the Contact Report. This creates a trust need for the AVACC:

- All data that is loaded into the AVACC computer systems, whether before launching or after, needs to be defined by the Mission Number.

a. Special Case of the Generate Contact Report (A.7.4.3)

The Contact Report is the sole means of communication between the AVACC and humans on the ship about the mission. Therefore, it must include all necessary information about the contact identification as well as any other information that is necessary to calibrate trust in the AVACC.

The Generate Contact Report function (A.7.4.3), as modeled, was not intended to be a Reasoning Function; the AVACC does not use automation to "decide" what is in, or how to format the Contact Report. The function is a summary output of the analysis performed by the AVACC.

The data provided in the Contact Report includes the Contact Characteristics as well as contact type, classification, and identification. As will be seen in the rest of the chapter, additional information that is deemed necessary for calibrating trust in the AVACC generated data also must be included in the Contact Report. The additional information is identified by analyzing the methods by which the AVACC completes Perform Mission functions for trust needs.

The Contact Report is the medium by which the AVACC and human interact; it is part of the interface. The contents of the Contact Report are how the AVACC's purpose, process and performance can influence trust calibration through knowledge, beliefs and perceptions. If the Generate Contact Report function was a Reasoning function, the AVACC would need to provide the "reasoning" behind its "decisions" to include or exclude data as part of the Contact Report. The "reasoning" would have to be understood and trusted by the humans in the interaction. Thus, an analysis into what attributes of the automation (purpose, process, and performance) can influence the human's trust (through knowledge, beliefs and perceptions) would be needed for this function. The Reasoning functions in this chapter are analyzed in this way, but Generate Contact Report is not. It is the author's opinion that increasing the number of automated functions in a system should be performed in an iterative manner. Before the AVACC can be given the authority to determine what needs to be communicated in the Contact Report, the contents of a Contact Report that communicates enough information to calibrate trust in the AVACC must be determined. Therefore, Generate Contact Report is a Non-Reasoning function.

The next trust need for the AVACC is therefore identified:

- The Contact Report needs to include sufficient information to calibrate human trust in the AVACC generated vessel identification, and the methods by which the identification was determined.

This trust need is obviously vague but can only be refined by analyzing the other Perform Mission functions for what "sufficient information" means.

2. Equation Functions

The next category, Equation Functions, includes six of the Non-Reasoning functions. As the name implies, these functions are completed by mathematically relating variables created by the sensors, time, and loaded Mission Data to each other to create new information about the contact. Each of the Equation function's results are included as part of the Contact Characteristics, and thus the Contact Report, and are used in Reasoning function algorithms as shown in the model in Chapter IV. The Equation Functions are listed in Table 3.

Table 3. Non-Reasoning Equation Functions

Equation Functions	
Number	Function Name
A.7.2.1.4	Determine Detection Range & Bearing from Self
A.7.3.3	Determine Contact Speed
A.7.3.4	Determine Contact Course
A.7.3.5	Determine CPA to HVU Critical Zone
A.7.3.6	Predict Contact Future Position
A.7.3.7	Predict Intercept to Contact

Together, sensor output and GPS data describe the position of a detection or contact. Similarly, GPS data provides the position of the AVACC. The HVU Critical Zones are defined by the Mission Data. With positions defined, and the elapsed time between subsequent positions recorded, all of the Equation functions can be calculated.

Equation functions that determine range (or distance) and bearing (or course) use trigonometry. Positions are defined by latitude and longitude from the GPS data, and changes in the latitude and longitude between two location points are first determined. The Pythagorean Theorem is applied to the latitude and longitude differences to determine range (or distance) between the two points. The inverse tangent function is

applied to the quotient of the change in longitude (the dividend) and the change in latitude (the divisor) to determine angular bearing (or course) from one point to another.

Functions that determine speed are calculated by dividing the distance between two subsequently recorded positions (determined by applying the Pythagorean Theorem) by the elapsed time between the positions. The two prediction functions determine future positions based on the contact current position, the calculated course, and the calculated speed of the entity of interest.

None of the Equation functions include any "reasoning" action. The results they provide are only affected by the accuracy of the variables in their respective calculations. The accuracy of the variables trace to the performance of the system sensors. The sensors' operational status can be provided to affect perception of the system. This was discussed in Chapter III, Section B. by paraphrasing the example studies from Parasuraman and Riley (1997, 236–237). Additionally, Lyons, summarizing Wang et al. (2013), agrees and states that "indicators of [sensor] reliability can be a useful piece of information to users as this may help the humans calibrate their trust" (Lyons 2013, 52). Equation function data can affect trust. Perception of the performance of the system sensors should be addressed to aid with calibrating trust in the AVACC. Therefore, another trust need is identified:

- System sensor operational statuses need to be communicated at the human-machine interface to affect proper calibration of trust. The operational status affects the perception of the AVACC's ability to meet its purpose and execute its processes.

3. Sensor Functions

The remaining nine Non-Reasoning functions fit into the Sensor Function category. The Sensor Functions are those that describe sensor actions and sensor specific data. They are listed, along with their associated sensors in Table 4.

Table 4. Non-Reasoning Sensor Functions

Sensor Functions		
Number	Function Name	Associated Sensor
A.7.1.1	Determine Self Position	GPS
A.7.1.3.3	Search Area with Sensors	Radar, ESM, Visual, IR
A.7.2.1.1	Define Detection Criteria	Radar, ESM, Visual, IR
A.7.2.1.2.1	Passively Detect EM Signal	ESM
A.7.2.1.2.2	Receive Radar Return	Radar
A.7.2.1.3	Geo-locate Contact Position	GPS, Radar, ESM, Visual, IR
A.7.4.1	Collect Contact Image	Visual, IR
A.7.4.4	Determine if Communication Established	Data Link (Transmitter/Receiver)
A.7.4.6	Transmit Contact Report	Data Link (Transmitter/Receiver)

The Sensor functions all involve operating one or more of the sensor subsystems installed on the AVACC. The purpose for each of the sensors subsystems needs to be addressed prior to interaction with the AVACC because they each help describe the capabilities and limitations of the system. Education and training are methods that can be used to affect knowledge and perception of the AVACC sensor's purposes and processes. The AVACC high-level capability needs such as operational range, power requirements, and size constraints to name a few, will determine the operational sensor characteristics. The sensor characteristics, based on the capability needs, set sensor thresholds. This information should be known prior to using the AVACC for an ISR mission. A baseline knowledge of the sensors' purposes and processes will affect how trust in the system is calibrated.

In addition to knowledge of sensor purposes and processes, system sensor performance must be addressed. Educating potential users on the AVACC system only provides a starting point for calibrating trust. As users interact with the AVACC via the Contact Report, perception of the system sensors' performance needs to be addressed. Perception of sensor performance is also identified as a need for the Equation functions.

C. REASONING FUNCTIONS

The Reasoning functions performed by the AVACC in the ISR mission include sorting, matching, and "deciding" actions. These type of functions are described throughout this investigation as those in which automation is used to mimic human thinking processes, or cognitive functions. The reader may have noticed that the ISR mission model in Chapter IV is developed so that it could be used to describe a manned system or an automated system; this was done intentionally. The model is a tool that the author uses to communicate the process steps that an aerial vehicle system must complete to identify a vessel on the open ocean to the reader. Even if the reader does not have experience with aerial vehicles or with ISR missions, the model can be followed and the process can be understood. Similarly, humans that interact with the AVACC may not have knowledge of, or experience with, how computer programs work. Therefore, it is necessary to consider what information about the processes must be supplied in the interaction to adequately describe how the AVACC's processes are completed.

In their study into judgment and trust factors in automated decision aids, Seong and Bisantz (2002, 247) found and state that "providing cognitive feedback information" to users at the human-automation interface "can actually allow human operators to understand the inner workings" of the automated system. In other words, providing information that describes how Reasoning functions are performed influences perception of system processes. However, care should be taken in how the cognitive feedback information (Seong and Bisantz 2002, 247), or in the author's terms, process explanation information, is provided. Lyons notes that "too much information, or non-intuitive displays may confuse and frustrate users" (2013, 52). If the AVACC's automation process explanation information is provided in the Contact Report but is perceived as confusing, trust will not be calibrated for appropriate reliance. Therefore, the correct information explaining the Reasoning function processes needs to be provided so that it can be used appropriately in the AVACC-human interaction.

Twelve of the Perform Mission functions are Reasoning functions, and they were further categorized by describing them, in general terms, by the assumed methods that the AVACC's computers would use to complete actions and produce information. These

categories are Optimization functions, Image Recognition and Image Matching functions, and Comparison functions. Each are described in more detail in the following sections.

1. Optimization Reasoning Functions

There are three Reasoning functions in the AVACC ISR mission model that are characterized as Optimization functions. They are listed in Table 5.

Table 5. Optimization Reasoning Functions

Optimization Functions	
Number	Function Name
A.7.1.3.2	Determine Optimal Search Pattern
A.7.3.10	Determine Best Resolution Location
A.7.4.5	Generate Maneuver Request

Each of the Optimization functions is performed by an algorithm, or series of algorithms written into the AVACC's computer systems. The purpose of each of the Optimization functions is to find a minimum, maximum, or pre-defined "best" solution to a question. For the identified functions, the questions are:

- For Determine Optimal Search Pattern, what is the "best" way to search a defined area given the limitations of the sensors installed on the AVACC, the AVACC location, the Mission Data limits (with included Intelligence Report), and the remaining mission time?
- For Determine Best Resolution Location, where is the best location, given the AVACC operational limits, sensor limits, Mission Data limits, and mission time to resolve a tracked contact?
- For Generate Maneuver Request, where should the AVACC position itself to reestablish communication with the ship based on current AVACC position, AVACC operating limitations, Mission Data, and last known position of the ship?

Arguably, the "best" solution for the three Optimizing functions could include minimizing time, maximizing search area coverage, maximizing sensor capabilities, and maximizing safe standoff distance from vessels, to name a few examples. The variables that are considered in the algorithms to reach any of these goals could each carry

different weighting factors depending on how the designer, or user, describes the “best” solution. If the designer’s view of “best” does not match the user’s view of “best,” or if the designer’s “best” is not clearly communicated to the user, trust calibration will be negatively affected. Users would be required to blindly accept (over trust) the AVACC’s methods, or they could potentially reject (under trust) the AVACC’s methods. Neither of these outcomes are desirable. In a trust calibrated interaction, the desire is to match the human’s reliance on the system with the capabilities of the system.

Prior knowledge of how the AVACC performs Optimization functions is necessary for trust calibration to achieve proper reliance on the system. However, because of the weighted nature of optimization processes, prior knowledge may not be enough to properly calibrate trust. Therefore, the Contact Report, which is the only method of communication in the modeled process, needs to include information that explains how optimization functions influence the identification of a vessel. The data from functions Determine Optimal Search Pattern (A.7.1.3.2) and Determine Best Resolution Location (A.7.3.10) would influence how contact identification is achieved. The AVACC Contact Report must include the required information for the process explanation to affect perception of AVACC process and also AVACC performance. The examples that the author listed for determining the “best” solution for the Optimization functions in general could all be factors that determine a search pattern. However, the elapsed mission time or other factors could potentially change what the “best” search pattern is at any given time. Why a search pattern was determined needs to be communicated, resulting in the following trust need:

- The Contact Report needs to include the search pattern that was determined by the AVACC and the determining factors for why it was used.

For similar reasons as those given for communicating search pattern determining factors, the resolution positon and factors that determined why the position was used need to be communicated. Therefore, another trust need is proposed:

- The AVACC generated Contact Report needs to include the Resolution Location, described by AVACC range and bearing from the contact and the determining factors that “decided” the Resolution Location.

2. Image Recognition and Matching Reasoning Functions

Five of the twelve Reasoning functions identified in the AVACC ISR mission model are characterized as Image Recognition and Image Matching functions. These are listed in Table 6.

Table 6. Image Recognition and Image Matching Reasoning Functions

Image Recognition and Matching Functions	
Number	Function Name
A.7.2.1.2.3	Detect a Visual Image
A.7.2.1.2.4	Detect an IR Image
A.7.4.2.2	Determine Contact Type
A.7.4.2.3	Determine Contact Classification
A.7.4.2.4	Determine Contact Identification

Imagery is the key component to all of the Image Recognition and Image Matching functions. For the AVACC (or a human) to identify a vessel at sea, it must be "seen." In Chapter IV, Section I, Google™'s image recognition software was mentioned as an example for how automation can be used to interpret visual information. However, the ability to interpret visual information is not sufficient unless how the interpretation was performed is understood and appropriately trusted.

For a human operator, the eye collects light waves, and the brain determines what the light waves mean. Similarly, the optics in visual sensors, and a combination of optics and a thermal sensor for IR sensors, collect EM energy. The sensor's processors together with the automation, determine what the collected EM energy means. The components are comparable in purpose, but not in process or performance.

Image recognition requires experience, knowledge, and perception. How humans put these together to decide (consciously or subconsciously) what is seen can functionally be described through models. Parasuraman et al. provide a simplified four stage model for how human information processing occurs. Decision Making is the third stage. They describe the first and second stages as the transition from sensory input and sensory processing through full "conscious perception and manipulation of processed and

retrieved information in working memory" (2000, 287) as prerequisites to decision making. A completely decomposed version of this model could be used as the foundation for writing algorithms to mimic the human process. However, the human decision-making model used has to be known and understood prior to, and during, the human-AVACC interaction.

Two notable, completely different methods are widely cited as models for how human decision making occurs, or should occur. The first method, referenced extensively by Danial Kahneman (2011) in his book "Thinking, Fast and Slow," was focused on the use of rigid algorithmic steps to reach conclusions. The argument for this method is that biases and heuristics can prevent humans from reaching the "best," or correct outcome for a given decision. Therefore, following defined procedural steps will yield the "best," or correct result.

However, depending on the variables used in the algorithmic steps, and how they are weighted, unexpected or irrational conclusions can occur. This is one reason why defenders of the second method for modeling, Naturalistic Decision Making (NDM), reject a strict algorithmic method (Kahneman 2011, 234) to describe human decision-making processes. Instead, NDM includes heuristics and biases in its methods. One NDM model, that could be used as a foundation for the AVACC Image Recognition and Matching functions is Klein's (1997, 285–292) Recognition Primed Decision Making (RPD) model, specifically the iteration of the model for expert decision making.

For automation to replace human decision making, it may make sense to model the automation processes after accepted models that describe an expert's method for decision-making. Dorman et al. (2016) provide an initial proof of concept for how RPD can be used as a foundation supervisory control of automation. One problem that arises in using RPD or other NDM methods for the AVACC is that a communication link is not continuous. The expert-based biases or heuristics would have to be programmed into the automation's algorithms and operate without supervision. This could lead to very complex algorithms that would be difficult to explain to a user; especially in the AVACC-human interaction. Further, using an expert's decision-making model as the

foundation requires programming expert-based “experience” into the AVACC and explaining to the humans how the “experience” affects the outcome.

Due to complexity issues, it seems that the strict algorithmic method must be followed for automation to complete the Image Recognition and Image Matching functions. However, the possibility of absurd or irrational results persists. Knowledge of the possibility that computer algorithms can error in this way is not enough to calibrate trust. Neither is knowledge of the method used in the automation’s “decision making.” Reasoning function results, specifically Image Recognition and Image Matching function results, must be explained during operational interactions to affect trust calibration in the AVACC, especially since methods cannot be observed without a constant communication link.

The algorithms used for the Image Recognition and Matching functions would interpret imagery and match elements in the image to known Mission Data elements that describe different vessel characteristics. The image elements that are used, and matched to the known elements, need to be communicated in the Contact Report.

- The AVACC Contact Report needs to include the image used to type, classify, and identify the vessel.
- The AVACC Contact Report needs to provide, and label the specific image characteristics used to determine the vessel type, classification, and identification.
- The AVACC Contact Report needs to include Mission Data elements that match contact image elements and are used for vessel type, classification, and identification.

3. Comparison Reasoning Functions

The remaining four Reasoning functions identified in the AVACC ISR mission model are characterized as Comparison functions. The Comparison functions are listed in Table 7.

Table 7. Comparison Reasoning Functions

Comparison Functions	
Number	Function Name
A.7.2.2	Assign Detection Type (Initial or Tracked)
A.7.3.2	Fuse Sensor Data/Contact Information
A.7.3.9	Determine if COI
A.7.4.2.1	Compare Collected Data to Known Information

Much like the Image Recognition and Matching functions, the Comparison functions relate collected sensor data to known information from the Mission Data loaded into the AVACC's computers prior to launch. The comparisons are completed by applying statistical methods in algorithms that again could follow different human decision-making models.

Assign Detection Type (A.7.2.2) determines if a sensor detection is an initial detection, which would result in developing a fix, or if a detection is associated with a tracked contact in the case where Detect (A.7.2) is a sub function of Track (A.7.3). The algorithms written for assigning the detection type would use probability-based methods to "decide" if the sensor data describes a new contact or not. Knowledge of the "decision" criteria used in the algorithm, an automation process, can affect the base line trust attitude toward the system but reporting the "decision" for all detections in the Contact Report is unnecessary and could confuse the user. This Reasoning function does not have a design specific trust need.

Fuse Sensor Data/Contact Information compares the data between different sensors. The algorithms written for this function could also use probability-based methods to determine if the data from the different sensors describe the same contact. If so, the data from each sensor is used to further refine Contact Characteristics such as location, course, and speed to name few. The Contact Characteristics are part of the Contact Report and knowing which sensor(s) were involved in developing characteristic data points can influence trust in the Contact Report information. Each of the sensors will have their own accuracy and precision, both of which can influence how well sensed data describes the same contact. Sensor accuracy and precision can influence the "decision"

process for or against combining two or more detections from different sensors together to describe a single entity. Knowing which data is Fused data, and why it is Fused, affects trust in the AVACC through perception of process, resulting in the following trust need:

- Contact Characteristics included in the Contact Report need to communicate the source of the characteristic data by the sensor(s) that created the data.
- Contact Characteristics that result from combining separate sensor data need to be identified in the Contact Report as Fused Data. The "reasoning" behind fusing the data needs to be communicated.

Determining if a contact is a COI requires comparing Contact Characteristics to predetermined COI trigger characteristics included in the Mission Data. Some example COI triggers could be contact distance from HVU Critical Zones, contact speed, contact course, or combinations of these individual characteristics to name a few. Much like the discussion for other functions, different variables that trigger a COI label could be weighted differently depending on how the algorithms or Mission Data are written. Why the AVACC "decided" to label a contact as a COI is important to the human's perception of the AVACC's process and how well it meets the purpose. This is captured in the following trust need.

- The Contact Characteristics that trigger labeling a contact as a COI needs to be included in the AVACC Contact Report.

The last of the Compare Functions, Compare Collected Data to Known Information, is procedurally right before Determine Contact Type in the ISR Process. All AVACC developed Contact Characteristics, with the exception of imagery, are compared to known Mission Data information to provide the Possible Contact List. The Possible Contact List directly influences and is refined by the type, classification, and identification determinations. Providing the entire Possible Contact List in the Contact Report could cause confusion; the list could be very large depending on the number of matches that exist between the Contact Characteristics and Mission Data. The entire list should not be provided in the Contact Report.

However, the Possible Contact list could unintentionally influence a contact identification. The Mission Data includes a loaded intelligence brief (or more accurately

intelligence based data points) and Ship Recognition Knowledge Database that may not include or adequately match what the AVACC "sees" in the image. In such cases the AVACC still needs to provide information about what has been found in the Contact Report. Therefore, if the AVACC cannot determine a vessel identification from the sensed and load data it has, the closest possible matches should be included in the Contact Report. Why any of the vessels on the Possible Contact List are "close" matches should also be communicated.

It is not the author's intent to state how many, or what Contact Characteristics or imagery elements, which could collectively be called Identification Factors, are needed to for the AVACC to complete a vessel identification. That determination would be made by the designers of the AVACC based on an analysis of the number of characteristics a human (possibly an expert) uses to complete a vessel identification. Instead, the intent is to ascertain what in the identification process needs to be communicated for a human to appropriately trust the automation generated solution. Prior to interacting with the AVACC, the human must know how many, and what types of Identification Factors are needed to make an identification. This establishes a baseline knowledge and perception of the AVACC's processes. In the interaction, the human perception of process and performance can be influenced to appropriately calibrate trust by communicating the number of and type of Identification Factors that are used in the identification or "close match" given in the Contact Report. This is captured in the following trust need:

- The Contact Report needs to list Identification Factors used in the identification "decision," and provide a total count for these factors.

D. REQUIREMENTS FROM TRUST NEEDS

Analyzing the Perform Mission functions of the AVACC ISR mission shows that trust needs do trace to the system; the needs do not only apply at the interface. This is an important point to consider in system design especially for systems that cannot be monitored during mission execution. The identified trust needs from this analysis can be used to develop AVACC software functional requirements. The trust needs and examples of requirements are provided in Table 8.

Table 8. AVACC Trust Needs and Example Requirements

AVACC Trust Needs	Associated Software Functional Requirements
All data that is loaded into the AVACC computer systems, whether before launching or after, needs to be defined by Mission Number.	➢ All data loaded into the AVACC computer systems shall include a Mission Number.
System sensor operational statuses need to be communicated at the human-machine interface to affect proper calibration of trust. The operational status affects the perception of the AVACC's ability to meet its purpose and execute its processes.	➢ The Contact Report shall include system statuses of all onboard subsystems.
The Contact Report needs to include the search pattern that was determined by the AVACC and the determining factors for why it was used.	➢ The AVACC search pattern route shall be included in the Contact Report. ➢ The factor(s) that the AVACC search pattern optimizes shall be included in the Contact Report.
The AVACC generated Contact Report needs to include the Resolution Location, described by AVACC range and bearing from the contact and the determining factors that "decided" the Resolution Location.	➢ The Resolution Location will be defined by bearing and range from the AVACC and shall be included in the Contact Report. ➢ The factor(s) that the Resolution Location optimizes shall be included in the Contact Report.
The AVACC Contact Report needs to include the image used to type, classify, and identify the vessel.	➢ The Contact Image used for determining contact type, classification, and identification shall be included in the Contact Report.
The AVACC Contact Report needs to provide, and label the specific image characteristics used to determine the vessel type, classification, and identification.	➢ Image characteristics used to determine contact type, classification, and identification shall be listed in the Contact Report.
The AVACC Contact Report needs to include Mission Data elements that match contact image elements and are used for vessel type, classification, and identification.	➢ Mission data matching image characteristics used to determine contact type, classification, and identification shall be listed in the Contact Report.

AVACC Trust Needs	Associated Software Functional Requirements
Contact Characteristics included in the Contact Report need to communicate the source of the characteristic data by the sensor(s) that created the data.	➢ All Contact Characteristics included in the Contact Report shall include the source(s) of the data as part of the characteristic data point.
Contact Characteristics that result from combining separate sensor data need to be identified in the Contact Report as Fused Data. The "reasoning" behind fusing the data needs to be communicated.	➢ Fused data included in the Contact Report shall include a "fused data" label as part of the characteristic data point. ➢ The rationale (trigger) that the AVACC's computers use to equate two or more different sensor's outputs as describing the same entity shall be listed with the respective fused data point.
The Contact Characteristics that trigger labeling a contact as a COI needs to be included in the AVACC Contact Report.	➢ The Contact Characteristic(s) that drive categorizing a contact as a COI shall be labeled as a COI Characteristic in the Contact report.
The Contact Report needs to list Identification Factors used in the identification "decision," and provide a total count for these factors.	➢ The Contact Report shall include the number, and nomenclature of positive Identification Factors used in the identification process.

VI. CONCLUSIONS AND RECOMMENDATIONS

The objective of this research was to develop trust need statements and system trust requirements that are traceable to functions in a process performed by a human-automation system. This objective was met by using established theory and empirical findings from multiple studies on trust and automation in the functional analysis of the AVACC on an ISR mission.

The functional analysis provided does not cover all process functions from AVACC launch to landing. Instead, the analysis focuses on the high-level Perform Mission function. For the AVACC to conduct an ISR without a communication link, the Perform Mission function must be completely allocated to the AVACC. Therefore, this function provides a good starting point for analyzing an "autonomous" system's functional architecture for trust needs.

A. DISCUSSION

The rationale for using a functional analysis as a method for identifying trust needs in a system is based on when the functional analysis occurs in the Systems Engineering (SE) process. Functional analysis occurs early in the SE process after the problem definition and high-level capability needs have been determined (Blanchard and Fabrycky 2011, 33–35). First, an operational scenario is developed to describe what actions must be completed to meet the capability needs and provide a solution to the problem. The actions, or functions, described in the scenario are then used to create the system's functional architecture; typically, the architecture is modeled-based. Architecture modeling methods such as FFBDs use a top down approach decomposing high-level functions into sub-functions to describe the sequential relationship of the functions. For the purposes of this research, the functional decomposition stops when a logical hierarchy of steps has been developed and the lowest level functions can be allocated to humans, hardware, or software components. At this point in the SE process, functions are analyzed and requirements are written for the system even though a specific design has not yet been determined. The trust needs analysis, and the development of

requirements based on those needs should occur at the same time. The requirements are the core of the system. System designers and testers ensure that the correct system is built to meet the defined needs by designing and testing to the written requirements.

For this research, an existing capability was chosen: the ability to perform an ISR mission without constant communication with a controlling station. The capability is currently met with a manned aerial system. However, recent developments in automation technology indicate that an advanced unmanned system, or combination of advanced systems, could feasibly be used instead of the manned system to meet the same capability.

If an advanced unmanned system, described in this research as the AVACC, were to replace a manned system for the ISR mission, the actions needed for mission accomplishment do not change very much. However, the functional allocation is completely different. Functions that were allocated to a human operator must now be assigned to software in the AVACC. In addition to function allocation, human-system interaction changes. With humans removed from the vehicle, the primary operational human-system interaction occurs when the AVACC provides a Contact Report to the humans on the ship. For the Contact Report to be properly accepted and used in any further ship mission decisions, the human must have calibrated trust in the AVACC and its ability meet the ISR capability. Thus, an investigation into trust in automation is needed to ensure the correct system is designed to meet capability needs.

The literature on trust and automation shows that human trust is calibrated through their knowledge, beliefs and perceptions of the automation's purpose, process, and performance (Lee and See, 2004; Madhavan and Weigmann, 2007; Ajzen, 2007). The functional analysis, specifically the development of the functional architecture, is the first place in the SE process that systems engineers can begin to analyze "autonomous systems" for trust needs. Performing trust analysis in conjunction with the functional analysis allows the systems engineer to develop trust requirements whose associated needs trace to the system's operational functions and factors that affect human trust.

The SE process is iterative. The analysis performed in this research is only one step towards realizing a total system architecture and set of requirements for designing an "autonomous" aerial system for ISR. This first step has provided an example for how to ensure human trust needs are addressed and designed into the next generation of automated systems. The foundation of a system is laid in the beginning of the SE process with the generation of requirements. If the foundation is weak, or if it is missing an important element, the entire system is doomed to fail.

B. RECOMMENDATIONS FOR FUTURE WORK

Expanding the trust analysis to a complete functional architecture would be the next step for continuing this research. Then, with trust needs identified and initial trust requirements written, measures of effectiveness and measures of performance used for testing the system for calibrated trust would be developed.

Unfortunately, clear measures of effectiveness and measures of performance for trust calibration cannot be adequately developed until an initial interface design is considered. The trust analysis performed in this work identifies what information must be provided by the AVACCC in the Contact Report based on human knowledge, beliefs and perception of the AVACC purpose, process and performance. It does not address how the information in the Contact Report will be put together at the human-AVACC interface.

From the trust analysis on the Perform Mission function, it is clear that the image of a vessel is critical to the identification process. The analysis also shows that providing the image at the interface can significantly aid in trust calibration. Therefore, it is the author's recommendation that the first iteration of the human-AVACC interface be a visual display that incorporates the vessel image with the other Contact Report information that can aid in trust calibration. The display arrangement, size, use of colors and shapes, as well other attributes will need to be determined by performing a human factors analysis on display design.

With an initial interface design in place, developmental testing for calibrated trust can begin. However, trust is not defined on an ordinal scale, nor is it directly measureable. As defined by the author in Chapter III of this research:

> Trust is the attitude of a human, developed from beliefs, perceptions and knowledge of a system's functional capabilities, towards the behavior of reliance in the system's actions to achieve the human defined goals in situations characterized by uncertainty and vulnerability.

The methods used in the majority of the empirical studies on trust and reliance in automation that are referenced in this research use reliance on automation as an indication of trust calibration. Desired trust calibration would be indicated by correct reliance on the AVACC generated, and interface displayed Contact Report information in multiple different situations with varying amounts of uncertainty and vulnerability. Once the initial developmental testing has occurred, the testing results would drive the next iteration of trust analysis and generation or refinement of trust requirements.

LIST OF REFERENCES

Ajzen, I. 2007. *Attitudes, Personality and Behavior*. McGraw-Hill Education. Accessed August 6, 2016. ProQuest Ebook Central. http://ebookcentral.proquest.com.libproxy.nps.edu/lib/ebook-nps/detail.action?docID=287791&fPQ=1

Bradshaw, Jeffery M., Robert R. Hoffman, Matthew Johnson, and David D. Woods. 2013. "The Seven Deadly Myths of "Autonomous Systems." *IEEE Intelligent Systems* 28(3): 54–61. Accessed July 24, 2016. doi 10.1109/MIS.2013.70.

Blanchard, Benjamin S. and Wolter J Fabrycky. 2011. *Systems Engineering and Analysis.* 5th ed. Upper Saddle River: Prentice Hall.

Defense Science Board. 2012. *Task Force Report: The Role of Autonomy in DOD Systems*. Department of Defense, Defense Science Board Office of the Under Secretary of Defense for Acquisition, Technology and Logistics. July 2012. Washington, DC, http://www.acq.osd.mil/dsb/reports/AutonomyReport.pdf.

———. 2016. *Report of the Defense Science Board Summer Study on Autonomy*. Department of Defense, Defense Science Board Office of the Under Secretary of Defense for Acquisition, Technology and Logistics. June 2016. Washington, DC, Accessed August 25, 2016. http://www.acq.osd.mil/dsb/reports/DSBSS15.pdf.

Department of Defense. 2012. *Autonomy in Weapons Systems.* Department of Defense Directive 3000.09. November 21, 2012. Deputy Secretary of Defense. Washington, DC, http://www.dtic.mil/whs/directives/corres/pdf/300009p.pdf.

———. 2015a. *Technical Assessment: Autonomy*. February 2015. Office of Technical Intelligence, Office of the Assistant Secretary of Defense for Research & Engineering. February 2015. Washington, DC, http://www.defenseinnovationmarketplace.mil/resources/OTI_TechnicalAssessment-AutonomyPublicRelease_vF.pdf.

———. 2015b. *Technology Investment Strategy 2015–2018*. Office of the Assistant Secretary Autonomy Community of Interest (COI) Test and Evaluation, Verification and Validation (TEVV) Working Group. May 2015. Washington, DC, http://www.defenseinnovationmarketplace.mil/resources/OSD_ATEVV_STRAT_DIST_A_SIGNED.pdf.

Dorton, Stephen, Brett Terry, Bobby Jaeger, and Peter B. Shearer. 2016. "Development of a Recognition Primed Decision Agent for Supervisory Control of Autonomy." *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situational Awareness and Decision Support (CogSIMA)*, 21–25 March 2016. Accessed October, 1, 2016. http://dx.doi.org/10.1109/COGSIMA.2016.7497806.

Frau, Pedro, Ricky Howell, Inez Kelly, Steve “K-9” Kulikowski, Hollen Mak, Tony Mikulin, Thang Nguyen, Nolan Paulsen, Beth Wade and Donna Young. 2011. “An Architecture For An Autonomous, Weaponized Unmanned Aerial System (UAS).” Technical Report, Naval Postgraduate School. NPS-SE-11-001.

Hernandez, Marc, Henry Jackson, Oscar Meza, Craig McKenny, Rebecca Morgan, Billy Palermo, Sommer Roach, Al Spaterna, and Diane Wathen. 2010. “MH-60 Seahawk / MQ-8 Fire Scout Interoperability.” Capstone Design Project, Naval Postgraduate School. NPS-SE-10-006.

Hoffman, Robert R., John D. Lee, David D. Woods, Nigel Shadbolt, Janet Miller, and Jeffery M. Bradshaw. 2009. “The Dynamics of Trust in Cyberdomains.” *IEEE Intelligent Systems*, Vol 24, Issue 6 (Nov-Dec): 5–11. Accessed July 24, 2016. http://dx.doi.org/10.1109/MIS.2009.124.

Hoffman, Robert R., Matthew Johnson, and Jeffery M. Bradshaw. 2013. “Trust in Automation.” *IEEE Intelligent Systems*, Vol. 28, Issue 1 (Jan-Feb): 84–88. Accessed July 24, 2016. http://dx.doi.org/10.1109/MIS.2013.24.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.

Klein, Gary. 1997. “The recognition-primed decision (RPD) model: Looking back, looking forward.” In *Naturalistic Decision Making*, edited by Caroline E. Zsambok and Gary Klein, 285–292. New Jersey: Lawrence Erlbaum Associates, Inc.

Klein, Gary, David D. Woods, Jeffery M. Bradshaw, Robert R. Hoffman, and Paul J. Feltovich. 2004. “Ten Challenges for Making Automation a ‘Team Player’ in Joint Human-Agent Activity.” *IEEE Intelligent Systems*, Vol 19, Issue 6 (Nov-Dec): 91–95. Accessed July 24, 2016. http://dx.doi.org/10.1109/MIS.2004.74.

Lee, John D. and Neville Moray. 1992. “Trust, control strategies and allocation of function in human-machine systems.” *Ergonomics*, Vol 35, No. 10: 1243–1270. Accessed July 28, 2016. http://dx.doi.org/10.1080/00140139208967392.

———. 1994. “Trust, self-confidence, and operators’ adaptation to automation.” *International Journal of Human-Computer Studies*, Vol 40, No. 1 (Jan):153–184. Accessed July 28, 2016. http://dx.doi.org/10.1006/ihhe.1994.1007.

Lee, John D. and Katrina A. See. 2004. “Trust in Automation: Designing for Appropriate Reliance.” *Human Factors*, Vol. 46, No. 1, Spring: 50–80. Accessed April 25, 2016. http://dx.doi.org/10.1518/hfes.46.1.50_30392.

Lyons, Joseph B. 2013. “Being Transparent about Transparency: A model for Human-Robot Interaction.” *AAAI Spring Symposium Series 2013*. Accessed September 26, 2016. http://www.aaai.org/ocs/index.php/SSS/SSS13/paper/view/5712.

Lockett III, John F. and Jeffrey Powers. 2003. “Human Factors Engineering Methods and Tools.” In *Handbook of Human System Integration*, edited by Harold R. Booher, 463–496. New Jersey: John Wiley and Sons, Inc.

Madhavan, P and D.A. Weigmann. 2007. “Similarities and differences between human-human and human-automation trust: an integrative review.” *Theoretical Issues in Ergonomics Science*, 8:4 (May): 277–301. Accessed April 25, 2016. http://dx.doi.org/10.1080/14639220500337708.

Masiello, Thomas J. Major General, USAF. 2013. *Air Force Laboratory Autonomy Science and Technology Strategy.* Air Force Research Laboratory. December 2013. Wright-Patterson AFB, Ohio. Accessed November 25, 2015. http://defenseinnovationmarketplace.mil/resources/AFRL_AutonomyStrategy-DistroA.pdf.

Muir, Bonnie M. 1988. “Trust between humans and machines, and the design of decision aides.” *International Journal of Man-Machine Studies*, Vol 27, Issues 5–6 (Nov, Dec): 527–573. Accessed July 28, 2016. http://dx.doi.org/10/10.16/50020-7373(87)80013-5.

Parasuraman, R., Victor Riley. 1997. “Humans and automation: Use, misuse, disuse, abuse.” *Human Factors*, Vol 39, Issue 2 (June): 230–253. Accessed July 26, 2016. http://dx.doi.org/10.1518/001872097778543886.

Parasuraman, Raja, Thomas B. Sheridan, and Christopher D. Wickens. 2000. “A Model for Types and Levels of Human Interaction with Automation.” *IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans*, Vol 4, No. 3 (May): 286–297. Accessed April 25, 2016. http://dx.doi.org/ 10.1109/ 3468.844354.

Proctor, Robert W. and Trisha Van Zandt. 2008. *Human Factors in Simple and Complex Systems.* 2nd ed. Boca Raton: CRC Press Taylor and Francis Group, LLC.

Riley, Victor. 1989. “A General Model of Mixed-Initiative Human-Machine Systems.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol 33, No. 2 (October): 124–128. Accessed April 25, 2016. http://dx.doi.org/ 10.1177/154193128903300227.

Seong, Younho and Ann M. Bisantz. 2002. “Judgement and Trust in Conjunction with Automated Decision Aids: A Theoretical Model and Empirical Investigation.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol 46, No. 3: 423–427. Accessed September 26, 2016. http://dx.doi.org/ 10.1177/154193120204600344.

Sheridan, Thomas B. 2002. *Humans and Automation: System Design and Research Issues*. Santa Monica, CA: Wiley & Sons, Inc.

Sheridan, Thomas B., and William L. Verplank. 1978. *Human and Computer Control of Undersea Teleoperators*. MIT Man-Machine Systems Laboratory. Technical Report 15 March 1977–14 June 1978. Cambridge, MA. Accessed August 7, 2016. http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA057655.

Wang, Lu, Greg A. Jameson, and Justin G. Hollands. 2009. "Trust and Reliance on an Automated Combat Identification System." *Human Factors*, Vol 51, No. 3 (June): 281–291. Accessed September 26, 2016. http://dx.doi.org/10.1177/0018720809338842.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California